

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).

Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

THIS PAGE LEFT BLANK

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

THIS PAGE LEFT BLANK

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 21:23:21 ; Search time 92.5884 Seconds
(without alignments)
4645.181 Million cell updates/sec

Title: US-10-728-323-1
Perfect score: 3586
Sequence: 1 aataatcatatattcatc.....cggttgctggtgtttctccc 2032

Scoring table:

BLOSUM62
Ygapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame_n2p.model -DEV=xlp
-Q=/abs/ABSSWEB_spool/US10728323/runat_15052006_172132_22377/app_query.fasta_1
-DB=UniProt -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=prco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US10728323 @CGN 1 1 694 @runat_15052006_172132_22377 -NCRPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3286	91.6	626	1 ALL12_ARAHY	P43238 arachis hyp
2	3286	91.6	626	2 Q547W5_ARAHY	Q547W5 arachis hyp
3	3041	84.8	614	1 ALL11_ARAHY	P43237 arachis hyp
4	2887.5	80.5	580	2 Q6PSU3_ARAHY	Q6PSU3 arachis hyp
5	2172.5	60.6	428	2 Q6PSU4_ARAHY	Q6PSU4 arachis hyp
6	1495.5	41.7	299	2 Q6PSU5_ARAHY	Q6PSU5 arachis hyp
7	1435	40.0	303	2 Q6PSU6_ARAHY	Q6PSU6 arachis hyp
8	1302	36.3	621	2 Q7XXT2_SOYBN	Q7XXT2 glycine max
9	1297	36.2	621	2 Q4LER6_SOYBN	Q4LER6 glycine max
10	1290	36.0	533	2 Q4LEB1_LUPAL	Q4LEB1 lupinus alb
11	1286	35.9	604	2 Q4LER5_SOYBN	Q4LER5 glycine max
12	1284	35.8	621	2 Q948Y0_SOYBN	Q948Y0 glycine max
13	1282	35.8	605	2 Q941X2_SOYBN	Q941X2 glycine max
14	1281	35.7	605	1 GLCA_SOYBN	P13916 glycine max
15	1267	35.3	571	1 CVCA_PEA	P13915 pisum sativ
16	1265.5	35.3	623	2 Q948X9_SOYBN	Q948X9 glycine max

17	1258.5	35.1	545	2	Q41674_VICNA	Q41674 vicia narbo
18	1251	34.9	613	1	Q9M3X6_PEA	Q9M3X6 pisum sativ
19	1251	34.9	639	1	GLCAP_SOYBN	P11827 glycine max
20	1190	33.2	559	2	Q9FZP9_SOYBN	Q9FZP9 glycine max
21	1183.5	33.0	418	2	Q84UI1_LENCU	Q84ui1 lens culina
22	1180.5	32.9	543	2	O22120_SOYBN	O22120 glycine max
23	1178.5	32.9	459	1	VCLC_PEA	P13918 pisum sativ
24	1168	32.6	415	2	Q84UI0_LENCU	Q84ui0 lens culina
25	1165	32.5	416	2	O22121_SOYBN	O22121 glycine max
26	1159	32.3	420	2	Q50JD8_SOYBN	Q50jd8 glycine max
27	1156	32.2	439	1	GLCB_SOYBN	P25974 glycine max
28	1152	32.1	439	2	Q93VL9_SOYBN	Q93vl9 glycine max
29	1148	32.0	438	2	Q43626_PEA	Q43626 pisum sativ
30	1145	31.9	415	2	Q702P1_PEA	Q702p1 pisum sativ
31	1135	31.7	415	2	O702P0_PEA	O702p0 pisum sativ
32	1123.5	31.3	463	2	Q41677_VICNA	Q41677 vicia narbo
33	1108.5	30.9	463	1	VCL_VICPA	P09438 vicia faba
34	1082	30.2	396	2	Q84UB3_SOYBN	Q84ub3 glycine max
35	1068	29.8	410	1	VCLB_PEA	P02854 pisum sativ
36	1030	28.7	518	2	Q9M3X8_LENCU	Q9M3X8 lens culina
37	962.5	26.8	445	1	CANA_CANGL	P50477 canavalia e
38	959.5	26.8	445	1	CANA_CANGL	P10562 canavalia g
39	910	25.4	593	2	Q9SEW4_PROSI	Q9sew4 juglans reg
40	907.5	25.3	810	2	Q92WI3_CUCMA	Q92wi3 cucurbita m
41	880	24.5	481	2	O7Y1C1_JUGNI	O7y1c1 juglans nig
42	865.5	24.1	666	2	Q9SPL4_MACIN	Q9spl4 macadamia i
43	863.5	24.1	625	2	Q9SPL3_MACIN	Q9spl3 macadamia i
44	849.5	23.7	666	2	Q9SPL5_MACIN	Q9spl5 macadamia i
45	839	23.4	605	1	VCLA_GOSHI	P09799 gossypium h

ALIGNMENTS

RESULT 1

ALL12_ARAHY

ID ALL12_ARAHY STANDARD; PRT; 626 AA.

AC P43238;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE Allergen Ara h 1, clone P41B precursor (Ara h 1).

OS Arachis hypogaea (Peanut).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;

OC Arachis.

OX NCBI_TaxID=3818;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=cv. Florunner;

RX MEDLINE=96013631; PubMed=7560062;

RA Burks A.W., Cockrell G., Stanley J.S., Helm R.M., Bannon G.A.;

RT "Recombinant peanut allergen Ara h I expression and IGE binding in

patients with peanut hypersensitivity.";

RL J. Clin. Invest. 96:1715-1721(1995).

RN [2]

RP CARBOHYDRATE-LINKAGE SITE ASN-521.

RX MEDLINE=20455243; PubMed=10998264; DOI=10.1006/abio.2000.4737;

RA Kolarich D., Altmann F.;

RT "N-glycan analysis by matrix-assisted laser desorption/ionization mass

spectrometry of electrophoretically separated nonmammalian proteins:

application to peanut allergen Ara h 1 and olive pollen allergen Ole e

1.";

RL Anal. Biochem. 285:64-75(2000).

CC -!- ALLERGEN: Causes an allergic reaction in human.

CC -!- SIMILARITY: Belongs to the 7S seed storage protein family.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

DR EMBL; L34402; AAB00861.1; -; mRNA.
DR HSP; P25974; IIPJ.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR Pfam; PF00190; Cupin.1; 2.
KW Allergen; Glycoprotein; Signal.
FT SIGNAL 1 25 Potential.
FT CHAIN 26 626 Allergen Ara h 1, clone P41B.
FT CARBOHYD 521 521 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 626 AA; 71345 MW; 1A6BBE4149D0E3 CRC64;

Alignment Scores:

Pred. No.: 1.66e-208 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.6% Indels: 0
DB: 1 Gaps: 0

US-10-728-323-1 (1-2032) x ALL12_ARAHY (1-626)

QY 50 ATGAGAGGAGGGTTTCTCCACTGATCTGTGTCTAGGGATCCCTGTCTGGCTTCAGTT 109
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGATGCAAGTATCATCTTACAGAAAGAAACAGAGAACCCCTGGCCAG 169
DB 21 SerAlaThrHisAlaLysSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGTGCCCTCAGAGTGTCTCAACAGGACCGATGACTTGAAGCAAAAGGCATCGAGTCT 229
DB 41 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGTGTCAACCAAGCTCGAGTATGATCTCTGTGTCTATGATCTCTGAGACACACTGGC 289
DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACACCAACCAACGTTCCCTCCAGGGAGGAGACAGTGGCCGCCAACCCGGAGACTAC 349
DB 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATCACCGCTCAACCCGAGAGAGAGAGGAGGCGCGATGGGACACAGCTGGACCG 409
DB 101 AspAspAspArgArgGlnProArgGluGluGlyGlyArgTyrProAlaGlyPro 120
QY 410 AGGAGCGCTGAAGAGAGAAGACTGGAGACAACCAAGAGAAGATTGGAGCGCAACAGT 469
DB 121 ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgProSer 140
QY 470 CATCAGACGCCAGGAATAAGCCCGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGluGluGluGluGlu 160
QY 530 CCAGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCTTCTACTTCCCGTCAAGG 589
DB 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCCGCTACGGGAACCAAAACGGTAGGATCCGGGTCTCGACAGGTTTGAC 649
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAGGTCAAGCGAGTTTCAGAAATCTCCAGATCACCGTATTTGGCGAGATCGAGGCCAAA 709
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTGTCTTCCCAAGCAGCTGATCTGATACATCTTGTATTATCCAGCAA 769
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTGACCGTAGCAAAATGGCAATTAACAGAAAGACTTTAATCTTGACGAG 829
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260

QY 830 GGCCATGCACCTCAGAAATCCATCCCGTTTCATTTCTCATCTTGAACCCCATGACAAC 889
DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACTCAGAGTAGTAAATCTCATGCCCGTTAAACACACCCCGCCAGTTTGGAGAT 949
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGGAGCAGCGAGACCAATCATCTTACTTGCAGGCGTTTACAGAGGAATACG 1009
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGAGCGCCCTTCAATCGGAATTCATAGATACGAGGCGTGTGTGTAGAGAGAAT 1069
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGAGAGAGCGAGCGATGAGTACTCGGAGTACTGAGAAC 1129
DB 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGGAGTGTATGATCAAAAGTGTCAAAGGACACGTTGAAGAACTTACTAAGCAGCT 1189
DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCGTCTCAAGAAAGCGCTCCGAGAGAGGAGATATACCAACCAACCAATCAACTG 1249
DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAGGCGAGCCCGATCTTTCTAAACAATTTGGGAAGTTATTGAGGTGAAGCCAGAC 1309
DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGACCCCGAGCTTCAAGCCTGAGCCTGGACATGATCTCATCTGTGTAGAGATCAAGAA 1369
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GAGCTTTCATGCTCCACACTTCAACTCAAGGCGATGTTATCTGTCGTCGCAACAAA 1429
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAlaAsnLys 460
QY 1430 GGAACCTGGAAACCTTGAACTCGTGGCTGTAAAGAAAAAGAGCAACAACAGAGGGGACGCGG 1489
DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnArgArg 480
QY 1490 GAAGAAGAGAGACGAGACGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1549
DB 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTGAGGAGAGCGGATGTGTTTCATATGCCAGCAGCTATCCAGTAGCCATC 1609
DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTGAAAAACAACACAGA 1669
DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTCAGGTGATAGGACCAATCTGATAGACCAAGAGAGAGAGAGAGAGAGAGAT 1729
DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTGGGTGAACAGTGTGAGAAGCTCATCAAAAACCAAGAGAGGAATCT 1789
DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTCTCGTCTCTCAATCTCAATCTCAATCTCCGCTCGTCTCTGAGAAAGAG 1849
DB 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerProGluLysGlu 600
QY 1850 TCTCTCAGAAAAGAGGATCAAGAGGAGGAGAAACCAAGGAGGAGGAGGAGGAGGAGGAG 1909
DB 601 SerProGluLysGluAspGlnGluGluGluGluGluGluGlyGlyGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTTAAC 1927

```
Db      621  IleLeuLysAlaPheAsn 626
RESULT 2
Q547W5_ ARAHY
ID   Q547W5_ARAHY PRELIMINARY;      PRT;   626 AA.
AC   Q547W5;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Major allergen Ara h1.
OS   Arachis hypogaea (Peanut).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC   Arachis.
OX   NCBI_TaxID=3818;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=22960226; PubMed=14597159; DOI=10.1016/j.molimm.2003.09.002;
RA   Viquez O.M., Konan K.N., Dodo H.W.;
RT   "Structure and organization of the genomic clone of a major peanut
RT   allergen gene, Ara h 1";
RL   Mol. Immunol. 40:565-571(2003).
DR   EMBL; AF432231; AAL27476.1; -; Genomic DNA.
SQ   SEQUENCE 626 AA; 71345 MW; 1A6BBB841490D0E3 CRC64;

Alignment Scores:
Pred. No.:      1,66e-208      Length:      626
Score:          3286.00        Matches:      626
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%    Mismatches: 0
Query Match:     91.6%          Indels:      0
DB:              2             Gaps:        0

US-10-728-323-1 (1-2032) x Q547W5_ARAHY (1-626)

QY      50  ATGAGAGGAGGGTTTCTCCACTGATGCTGTGCTAGGAGTCTGTCTCTGGCTTCAGTT 109
Db      1  MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyLeuValLeuAlaSerVal 20
QY      110 TCTGCAACCATGCGCAAGTCATCACCCTTACGAGAGAAACAGAGAACCCCTCGCGCCAG 169
Db      21  SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY      170 AGTGCTCCACAGATTGTCAACAGGAACCGGATGCTTCAAGCAAAAGGCATGCCAGTCT 229
Db      41  ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysAlaCysGluSer 60
QY      230 CGCTGCACCAAGCTCAGATGATGATCCTCGTTGTGCTATGATCCTCGAGGACACACTGGC 289
Db      61  ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY      290 ACCACCAACACAGTTTCCCTCCAGGGAGCGGACACGTCGGCGCCAAACCCGAGACTAC 349
Db      81  ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY      350 GATGATACCCGCTCAACCCCGAGAGAGAGGAGGCGCGATGGGACCACTGGACCG 409
Db      101  AspAspAspArgArgGlnProArgGlnGluGlyArgTyrGlyProAlaGlyPro 120
QY      410 AGGGAGCGGTGAAGAGAAAGAACTGGAGACAAACCAAGAGAAGATTGGAGGCGCACCAAGT 469
Db      121  ArgGluArgGluArgGluGluAspTyrArgGlnProArgGluAspTrpArgArgProSer 140
QY      470 CATAGCAGCCACGGAAAATAAGCCCGAAGAGAGAGAGAAACAGAGTGGGAAACA 529
Db      141  HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGlnIleArgGlyThr 160
QY      530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGACCAACCCCTTCTACTTCCGCTCAGG 589
Db      161  ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180

QY      590 CGGTTTAGCACCCCGTACGGGAACCAAAACGGTAGATCCCGGTCTCTGCAGAGGTTTGAC 649
Db      181  ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY      650 CAAAGGTCAAGGACGTTTCAGAAATCTCCAGAAATCACCTATTGTCCAGATCCAGGCCAAA 709
Db      201  GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY      710 CCTAACACTCTTGTCTTCTCCCAAGCAGCGTGTGTGATAACATCTCTTATTCACAGCAA 769
Db      221  ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY      770 GGGGAAGCCACCGTACCGTACCAATATGCAATACAGAAAGAGCTTTAATCTTGACGAG 829
Db      241  GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY      830 GGCCATGCACACTCAGAAATCCCATCCCGTTTCATTTCTTACATCTTGAACCGCCATGCAAC 889
Db      261  GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY      890 CAGAACCTCAGAGTAGTAAATCTCCATGCCGTTTAAACACACCCGCGCAGTTTGAGGAT 949
Db      281  GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY      950 TTCTTCCCGGAGCAGCGCAGCAATCATCTTCTTTCACGGGCTTCAGCAGGAATACG 1009
Db      301  PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY      1010 TTGGAGGCGGCTTCAATCCGGAATTCATGATGATACGGAGGCTCTGTAGAGAGAT 1069
Db      321  LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn 340
QY      1070 GCAGGAGGTGAGCAAGAGGAGAGAGCGGATGAGTACTCGGAGTAGTAGTGAAGAAC 1129
Db      341  AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY      1130 AATGAAGAGTATAGTCAAGTGTCAAGGAGCAGCTTGAAGAACTTACTAAGCAGCT 1189
Db      361  AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY      1190 AAATCCGTCTCAAAAGAGGCTCCCAAGAGAGGAGATATCAACCAACCAATCACTTG 1249
Db      381  LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY      1250 ACAGAAAGCGAGCCCGCTCTTCTTAACTTTGGGAAGTATTATTGAGGTGAAGCCAGAC 1309
Db      401  ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY      1310 AAGAAGAACCCCGCTTCAGGACCTGCACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
Db      421  LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY      1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGGTTATCGTCGTCGTCAACAAA 1429
Db      441  GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
QY      1430 GGAACCTGAAACCTTGAACCTCGTGTGTGAAGAAAGCAACAACAGAGGGGACGGCGG 1489
Db      461  GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnArgGlyArgArg 480
QY      1490 GAAGAAGAGGAGGACGAAGACGAAGAGAGGAGGAAGTAAACAGAGAGGTGCGTAGGTAC 1549
Db      481  GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY      1550 ACAGCGAGGTTGAAGGAAGGCGATGTGTTTCATCTCCAGCAGGCTCATTCAGTAGCCATC 1609
Db      501  ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY      1610 AACGCTTCTCCGAACTCCATCTGCTTGGCTTCGATCAACCGCTGAAACCAACACACAGA 1669
Db      521  AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY      1670 ATCTTCTTGCAGGTGATAAGGACAATGTGTATAGACCAATAGAGAAAGCAAGCAGGAT 1729
```

Db 541 IlePheLeuAlaGlyAspLysAspValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATCCCTCGGGTCGGGTGAACAGTGTGAGAGCTCATCAAAACAGAGGAATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTGTCTGCTCTCAATCTCAATCTCAATCTCCGTCGTCCTCTGAGAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerProGluLysGlu 600
QY 1850 TCTCCTGAGAAAGAGCATCAAGAGGAGGAGAAACCAAGGAGGAGGAGGTCCTCTTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626
RESULT 3
ALL11 ARAHY STANDARD; PRT; 614 AA.
ID ALL11 ARAHY STANDARD; PRT; 614 AA.
AC P43237;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Allergen Ara h 1, clone p17 precursor (Ara h I).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC Rosids; eutrosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=cv. Florunner;
RC MEDLINE=96013631; PubMed=7560062;
RX Burks A.W., Cockrell G., Stanley J.S., Helm R.M., Bannon G.A.;
RA "Recombinant peanut allergen Ara h I expression and IGE binding in
RT patients with peanut hypersensitivity.",
RT J. Clin. Invest. 96:1715-1721(1995).
RN [2]
RP CARBOHYDRATE-LINKAGE SITE ASN-516.
RX MEDLINE=2045243; PubMed=10998284; DOI=10.1006/abio.2000.4737;
RA Kolarich D., Altman F.;
RT "N-glycan analysis by matrix-assisted laser desorption/ionization mass
RT spectrometry of electrophoretically separated nonmammalian proteins:
RT application to peanut allergen Ara h 1 and olive pollen allergen Ole e
RT 1.",
RL Anal. Biochem. 285:64-75(2000).
CC -1- ALLERGEN: Causes an allergic reaction in human.
CC -1- SIMILARITY: Belongs to the 7S seed storage protein family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; L38853; AAA60336.1; -; mRNA.
DR HSSP; P25974; 1LPJ.
DR IntAct; P43237; -.
DR GlycoSuiteDB; P43237; -.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR Pfam; PF00190; Cupin.1; 2.
KW Allergen; Glycoprotein; Signal.
FT SIGNAL 1 25 Potential.
FT CHAIN 26 614 Allergen Ara h 1, clone P17.
FT CARBOHYD 516 516 N-linked (GlcNAc...)/FTID=CAR_000218.
FT SEQUENCE 614 AA; 70283 MW; 1DDACF2I7ECSF31 CRC64;

Alignment Scores: 2,57e-192 Length: 614
Pred. No.: 3041.00 Matches: 595
Score: 96.0% Conservative: 9
Percent Similarity: 94.6% Mismatches: 7
Best Local Similarity: 84.8% Indels: 18
Query Match: 1 Gaps: 7
DB: 1
US-10-728-323-1 (1-2032) x ALL11_ARAHY (1-614)
QY 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTGTGTAGGATCCTTGTCTCGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGCATGCCAAGTCATCCTTACAGAGAGAAACAGAGAACCCCTCGCCACG 169
Db 21 SerAlaThrGlnAlaLys---SerProTyr---ArgLysThrGluAsnProCysAlaGln 38
QY 170 AGGTGCTCCAGAGTTGTCAACAGGAAACCGGATGACTTGAACAAAGGATCGCAGTCT 229
Db 39 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 58
QY 230 CGCTGCACCAAGCTCAGATATCATCTCGTTGTCTATGATCTCGAGGACACACTGCG 289
Db 59 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly 74
QY 290 ACACCAACCAAGTTCCTCCAGGGAGCGGACACGTCGGCGGCCCAACCCGAGACTAC 349
Db 75 AlaThrAsnGlnArgHisProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 94
QY 350 GATGATGACCGCGTCAACCCCAAGAGAGAGGAGGCGGATGGGACCCAGCTGACCG 409
Db 95 AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGluPro 114
QY 410 AGGAGCGTGAAGAGAGAAAGACATCTCGGAAACCAACCTTCTTACTTCCCGTCAAGG 589
Db 115 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgProSer 134
QY 470 CATCAGCAGCCAGGAAATAAGGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Db 135 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGlyTrpGlyThr 154
QY 530 CAGGTAGCCATGTGAGGAGAGAAACATCTCGGAAACCAACCTTCTTACTTCCCGTCAAGG 589
Db 155 ProGlySerGluValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 174
QY 590 CGGTTTAGCACCCTGACCGGAAACCAAAACCGTAGGATCCGGTCTCTGAGAGTTTGAC 649
Db 175 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 194
QY 650 CAAAGTCAAGGCAGTTTCAGATCTCCAGAAATCACCGTATTCTGCAGATCCGAGGCCAA 709
Db 195 GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg 214
QY 710 CCTAACACTCTTGTCTTCTCCCAAGCACGCTGATGCTGTATAACATCTTGTATCCAGCAA 769
Db 215 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 234
QY 770 GGGCAAGCCACCGTACCGTACCAATGGCAATAACAGAAAGAGCTTTAATCTTGTAGCAG 829
Db 235 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 254
QY 830 GGCATGCACTCAGAAATCCATCCGTTTCATTTTCTTACATCTTGAACCCCATGACAAC 889
Db 255 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 274
QY 890 CAGAACCTCAGAGTAGCTAAATCTCATGCCCGTTTAAACACACCCCGCCAGTTTGTAGGAT 949
Db 275 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 294
QY 950 TTCTTCCCGGAGCAGCGAGACCAATCATCTTCTGAGGGCTTCAGCAGGAATACG 1009
Db 295 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 314

QY	650	CAAAGGTCAAGCGAGTTTCAGAAATCTCCAGAATCAACCGTATTGTGCAGATCGAGGCCAAA	709
Db	195	GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg	214
QY	710	CCTAACACTCTTGTCTTCCCAAGCAGCGTGTAGTAAACATCTTGTGTATCCAGCAA	769
Db	215	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	234
QY	770	GGCAAGCCACCGTAGCCGTAGCAAAATGCGCAATACAGAAAGAGCTTTAATCTTCAGCAG	829
Db	235	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	254
QY	830	GGCCATCGACTCAGAAATCCCATCCGTTTCATTTCTACATCTTGAAACGCCCATGACAAAC	889
Db	255	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	274
QY	890	CAGAACTCAGACTAGCTAAATCTCCATGCCCGTTAAACACACCCGCCAGTTTGAGGAT	949
Db	275	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	294
QY	950	TTCTTCCCGCAGCAGCCGAGACCAATCATCTACTTGCAGGCGTTTCAGCAGGAATACG	1009
Db	295	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	314
QY	1010	TTGAGGCGCCCTTCAATGCGGAATTCATGAGATACGAGGCGTCTGTGTAGAGAGAAT	1069
Db	315	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn	334
QY	1070	GCAGGAGGTGAGCAAGGAGGAGCGGCGAGCGATGAGTACTCGGAGTAGTGAGAAC	1129
Db	335	AlaGlyGlyGluGlnGluArgGlyGlnArgArgSerThrArgSerSerAsp---	353
QY	1130	AATGAAGGAGTGATGCTCAAAGTGTCAAAGGACACGTTGAAGAACTTACTAAGCAGCT	1189
Db	354	AsnGluGlyValIleValLysValSerLysGluHisValGlnGluLeuThrLysHisAla	373
QY	1190	AAATCCGTCTCAAGAAAGCTCCGAAGAAGAGGGAGATATCAACCAACCAATCAACTTG	1249
Db	374	LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu	392
QY	1250	AGAGAAGCGAGCCCGACTTCTTCAACACTTTGGGAAGTTATTTCAGGTGAAGCCAGAC	1309
Db	393	ArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLysProAsp	412
QY	1310	AAGAAGAACCCCGAGCTTCAGGACCTGGACATGATGCTCACTGTGTAGAGATCAAAGAA	1369
Db	413	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu	432
QY	1370	GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCGTCGCTGCTCAACAAA	1429
Db	433	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys	452
QY	1430	GGAACCTGGAACCTTGAACCTCGTGGCTGTAAAGAAAGAGCAACAAACAGAGGGGACGCGG	1489
Db	453	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg	472
QY	1490	GAA-----GAAAGGAGGAGCGAAGACGAAGAAGAGGAGGGAAGTAAACAGAGAGGTG	1540
Db	473	GluGlnGlnLrpGluGluGluGluGluAspGluGluGluGlySerAsnArgGluVal	492
QY	1541	CGTAGGTACACAGGAGTTGAAGGAAGGCGATGTGTTTCATCATGCCAGCAGCTCATCCA	1600
Db	493	ArgArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisPro	512
QY	1601	GTAGCCATCAACGCTTCTCCGAACCTCACTCTGCTTGGCTTCGGTATCAACGCTGAAAC	1660
Db	513	ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn	532
QY	1661	AACCACAGAAATCTTCTTGAGGTGATPAGGACAAATGTGTATGACCCAGATAGAGAACAA	1720
Db	533	AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln	552


```
QY 1604 GCCATCAAGCTTCTCCGCAACTCCATCTGTTGGCTTCGGTATCAACGCTGAAAAACAC 1663
Db      ::::::::::::::::::::::
531 ValValAsnAlaThrSerAspLeuAsnPhePheAlaPheGlyLeuAsnAlaGluAsnAsn 550
QY 1664 CACAGAAATCTTCCTGCAGGTGATAGGACAATGTGATAGACAGATAGAGAAGCAAGCG 1723
Db      |||
551 GlnArgAsnPheLeuAlaGlySerLysAspAsnValIleSerGlnIleProSerGlnVal 570
QY 1724 AAGGATTTAGCATCTCCCTCGGTGCGGTGAACAAGTTGAGAAGCTCATCAAAAACCAAGAG 1783
Db      ::::::::::::::::::::::
571 GlnGluLeuAlaPheProGlySerAlaLysAspIleGluAsnLeuIleLysSerGlnSer 590
QY 1784 GAATCTCACTTTCTGAGTCTGCTCTCAATCTCAATCTCAATCTCCGTCGTCCTCGAG 1843
Db      |||
591 GluSerTyrPheValAspAlaGlnProGln----- 600
QY 1844 AAGAGTCTCTGAGAAAGAGATCAAGAGGAGGAAAAACCAAGGAGGGAAGGTGCATCTC 1903
Db      ::::::::::::::
601 -----GlnLysGluGluGlyAsnLysGlyArgLysGlyProLeu 613
QY 1904 CTTTCAATTTGAAGGCTTTT 1924
Db      |||
614 SerSerIleLeuArgAlaPhe 620
```

RESULT 10

```
Q6EBC1_LUPAL
ID Q6EBC1_LUPAL PRELIMINARY; PRT; 533 AA.
AC Q6EBC1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Beta-conglutin.
OS Lupinus albus (White lupin).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosoids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.
RX NCBI_TaxID=3870;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RA Monteiro S.A., Freitas R.M., Teixeira A.N., Ferreira R.B.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY500372; AAS97865.1; -; mRNA.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin_region.
DR Pfam; PF00190; Cupin_1; 2.
SQ SEQUENCE 533 AA; 62129 MW; F17B49A6F2C4B9EC CRC64;
```

Alignment Scores:

```
Pred. No.: 1.29e-76 Length: 533
Score: 1290.00 Matches: 273
Percent Similarity: 62.9% Conservative: 117
Best Local Similarity: 44.0% Mismatches: 132
Query Match: 36.0% Indels: 98
DB: Gaps: 17
```

US-10-728-323-1 (1-2032) x Q6EBC1_LUPAL (1-533)

```
QY 44 GGAGCAATCAGAGGGAGGTTTCTCCATGATCTGTTCTCTAGGGATCTTGTCTCTGGCT 103
Db      |||
2 GlyLysMetArgValArgPheProThrLeuValLeuValLeuGlyValPheLeuMet 21
QY 104 TCAGTTTCTGCACGATGCCAGTATCATCTTACCAGAAGAAACACAGAACCCCTGC 163
Db      ::::::::::::::
22 AlaValSerIleGlyIleAla-----TyrGlyGluLysAsp----- 33
QY 164 GCCCAGAGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATGC 223
Db      ::
34 -----ValLeuLysSerHisGluArgProGluGluArgGluGluGln----- 47
QY 224 GAGTCTCGTCACCAAGCTCGAGTATGATCTCGTTGTGTATGATCTTCGAGACAC 283
Db      |||
```

```
Db      -----GluTrpGlnProArg----- 52
QY 284 ACTGSCACCAACCAACAGTTCCCTCCAGGGAGCGACACGTGGCCGCAACCCGGA 343
Db      |||
53 -----ArgGlnArgPro----- 56
QY 344 GACTACGATGATACCGCGCTCAACCCCAAGAGAGAGGAGGAGCGCATGGGACCAAGCT 403
Db      |||
57 -----GlnSerArgArgGluArgGluGlnGluGlnGluGlnGlnGlnGln 73
QY 404 GGACCGAGGAGCGTGAAGAGAAAGAACAGCTGAGACAAACCAAGAGAAAGATTGGAGGCCA 463
Db      |||
74 TyrProArgArgGlnSerGlyTyrGlu-----ArgArg 84
QY 464 CCAAGTCTATCAGCAGCCAGGAAATAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGTGG 523
Db      |||
85 GlnTyrHisGlu-----ArgSerGluGlnArgGluGluArgGluGlnGlnGlnGln 102
QY 524 GGAACACAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCTTTCTTACTTCCCG 583
Db      |||
103 GlySerPro---SerTyrSerArgArgGln-----ArgAsnProTyrHisPheSer 118
QY 584 TCAAGCGGTTTAGCACCGCTACCGGAACCAAAACGGTAGGATCCGGTCTCTGCAGAGG 643
Db      |||
119 SerGlnArgPheGlnThrLeuTyrLysAsnArgAsnGlyLysIleArgValLeuGluArg 138
QY 644 TTTGACCAAAAGTCAAGGCAGTTTCCAGAATCTCCAGAATCCACCGTATTGTGCAGATCGAG 703
Db      |||
139 PheAspGlnArgThrAsnArgLeuGluAsnLeuGlnAsnTyrArgIleValGluPheGln 158
QY 704 GCCAAACCTTAACACTCTTGTCTTCCCAAGCACGCTGATGTGTGATAACATCTTGTATTC 763
Db      |||
159 SerLysProAsnThrLeuIleLeuProLysHisSerAspAlaAspTyrValLeuValVal 178
QY 764 CAGCAAGGGCAAGCCACGTCACCTAGCAATGCAATGCAATACAGAAAGAGCTTTAATCTT 823
Db      |||
179 LeuAsnGlyArgAlaThrIleThrIleValAsnProAspArgArgGlnAlaTyrAsnLeu 198
QY 824 GACGAGGGCCATGCACCTCAGAAATCCATCCGTTTTCATTCTCATCTTTGAACGCCAT 883
Db      |||
199 GluTyrGlyAspAlaLeuArgIleProAlaGlySerThrSerTyrIleLeuAsnProAsp 218
QY 884 GACAAACAGAACCTCAGAGTAGTATAAATCTCCATGCCCGTTTAAACACACCCGCCAGTTT 943
Db      |||
219 AspAsnGlnLysLeuArgValValLysLeuAlaIleProIleAsnAsnProGlyTyrPhe 238
QY 944 GAGGATTTCTCCCGCGAGCAGCCGACCAATCATCTCTACTTCGAGGGCTTCAGCAGG 1003
Db      |||
239 TyrAspPheTyrProSerSerThrLysAspGlnGlnSerTyrPheSerGlyPheSerArg 258
QY 1004 AATCGTTTGGAGGCGCTTCAATCGGAATTTCAATGAGATACGAGGAGTGTGTTA--- 1060
Db      |||
259 AsnThrLeuGluAlaThrPheAsnThrArgTyrGluGluIleGlnArgIleLeuGly 278
QY 1061 ---GAAGAGAATGCAGAGGAGTGCAGCAAGAGAGAGGAGCGAGGAGGAGTGGAGTACG 1117
Db      |||
279 AsnGluAspGluGlnGluTyrGluGluGlnArgArgGlyGlnGluGln----- 294
QY 1118 AGTAGTGACAAATGAAGGAGTGTAGTCAAAAGTGTCAAGGAGCAGCGTTGAAGAACTT 1177
Db      |||
295 ---SerAspGlnAspGluGlyValIleValSerLysLysLeuGlnIleGlnLysLeu 313
QY 1178 ACTAAGCAGCTAAATCCGCTCAAAAGAAAGGCTCCGAAGAGAGAGGAGATATACCAAC 1237
Db      |||
314 ThrLysHisAlaGlnSerSerSer-----GlyLysAspLysProSerAsp---SerGly 330
QY 1238 CCAATCAACTTGAGAGAAGCGGAGCCGATCTTCTTAACAACACTTTGGGAAGTTATTGAG 1297
Db      |||
331 ProPheAsnLeuArgSerAsnGluProIleTyrSerAsnLysTyrGlyAsnPheTyrGlu 350
QY 1298 GTGAGCCAGACAGACAAGAACCCAGCTTCAGGACCTGGACATGATGCTACCTGCTGTA 1357
Db      |||
351 IleThrProAsp---ArgAsnProGlnValGlnAspLeuAsnIleSerLeuThrTyrIle 369
```



```
QY 623 AGGATCCGGTCTCGAGAGGTTTGACCAAGGTCGAAGGAGTTTCAGAAAT 682
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
229 HisValArgValLeuGlnArgPheAsnLysArgSerGlnGlnLeuGlnAsnLeuArgAsp 248
QY 683 CACCGTATTGTGAGATCGAGGCCCAACCTTACACTCTTCTCTCCCAAGCAGCTGAT 742
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
249 TyrArgGluLeuGluPheAsnSerLysProAsnThrLeuLeuLeuProHisAlaAsp 268
QY 743 GCTGATAACATCTTGTATTACAGCAAGGCCAGCCCGTCAGCCGTAGCAATGCAAT 802
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
269 AlaAspTyrLeuLeuValLeuLeuAsnGlyThrAlaLeuLeuThrLeuValAsnAsp 288
QY 803 AACAGAAAGAGCTTTAATCTTGACGAGGCCATGCACTCAGAAATCCATCCGCTTTCATT 862
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
289 AspArgAspSerTyrAsnLeuGlnSerGlyAspAlaLeuArgValProAlaGlyThrThr 308
QY 863 TCCTACATCTTGAAACCGCCATGACACCAACCACTCAGATAGCTAAATCTCCATGCC 922
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
309 TyrTyrValValAsnProAspAsnAspGluAsnLeuArgMetIleThrLeuAlaIlePro 328
QY 923 GTTAACACACCCCGCCAGTTTGAGGATTTCTCCCGCGAGCAGCCGAGCAATCATCC 982
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
329 ValAsnLysProGlyArgPheGluSerPheLeuSerSerThrGlnAlaGlnSer 348
QY 983 TACTGAGGGCTTACAGCAAGATACGTTGGAGCGCCTTCAATCGGAAATCAATGAG 1042
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
349 TyrLeuGlnGlyPheSerLysAsnIleLeuGluAlaSerTyrAspThrLysPheGluGlu 368
QY 1043 ATACGAGGGTGTGTTAGAGAAGATGCGAGAGGTGAGCAA--GAGGAGCAGGCGCAG 1099
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
369 IleAsnLysValLeuPheGlyArgGluGluGlyGlnGlnGlyGluGluArgLeuGln 388
QY 1100 AGCGGATGGAGTACTCGGAGTAGTGAGACAAATGAAGAGTAGTAGTCAAGTGTCAAAG 1159
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
389 -----GluSerValIleValGluIleSerLys 397
QY 1160 GAGCAGCTTGAAGACTTACTAAGCAGCGTAAATCCGTCTCAAGAAAGGCTCCGAGAA 1219
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
398 LysGlnIleArgGluSerLysArgAlaLysSerSerArgLysThrIleSerSer 417
QY 1220 GAGGAGATATACCAACCAATCACTTGAGAGAAGCGGAGCCGATCTTCTCAACAAC 1279
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
418 Glu-----AspLysProPheAsnLeuArgSerArgAspProIleTyrSerAsnLys 434
QY 1280 TTTGGGAAGTATTATTGAGTGAAGCCAGACAAGAACCCCGAGCTTCAGGACCTGGAC 1339
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
435 LeuGlyLysLeuPheGluIleThrProGlu--LysAsnProGlnLeuArgAspLeuAsp 453
QY 1340 ATGATGCTACCTGTGTAGAGATCAAGAAGGAGCTTTGATGCTCCACACTTCAACTCA 1399
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
454 ValPheLeuSerValValAspMetAsnGluGlyAlaLeuPheLeuProHisPheAsnSer 473
QY 1400 AAGCCATGTTATCGTCGTCGTCACAAAGGAACCTGGAACCTTGAACCTCGGCTGTA 1459
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
474 LysAlaIleValValLeuValIleAsnGluGlyGluAlaAsnIleGluLeuValGlyIle 493
QY 1460 AGAAAAGAGCAACAACAGAGGGGACCGCGGGAAGAAGAGGAGGACGGAAGAGAGAG 1519
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
494 ---LysGluGlnGlnArgGlnGlnGlnGluGlnPro----- 506
QY 1520 GAGGAAGTAACAGAGAGTGTGCTAGGTACACAGCGAGGTTGAAGNAGCGGATGTGTC 1579
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
507 -----LeuGluValArgLysTyrArgAlaGluLeuSerGluGlnAspIlePhe 522
QY 1580 ATCATGCCAGAGCTCATCAGTAGCCATCAACGCTTCTCCCGAATCCCATCTGCTTGGC 1639
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
523 ValIleProAlaGlyTyrProValValValAsnAlaThrSerAspLeuAsnPhePheAla 542
QY 1640 TTCGGTATCAACGCTGAAAAACAACACAGAAATCTTCTTCAGGTGATAGGACAATGTG 1699
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
543 PheGlyIleAsnAlaGluAsnAsnGlnArgAsnPheLeuAlaGlySerLysAspAsnVal 562
```

```
QY 1700 ATAGACCAGATAGAGCAAGCGAGGATTTAGCATTCCTCGGTGCGGTGAACAAGTT 1759
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
563 IleSerGlnIleProSerGlnValGlnGluLeuAlaPheLeuGlySerAlaLysAspIle 582
QY 1760 GAGAAGCTCATCAAAACCCAGAAAGAAATCTCACTTTGTGAGTGTCTCAATCTCAA 1819
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
583 GluAsnLeuIleLysSerGlnSerGluSerTyrPheValAspAlaGlnProGln----- 600
QY 1820 TCTCAATCTCCGCTGCTCTCTGAGAAAGAGTCTCTCTGAGAAAGAGGATCAAGAGGAG 1879
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
601 -----GlnLysGluGluGly 605
QY 1880 AACCAAGGAGGAGGTCACCTCTTTCATTTGAAGGCTTTT 1924
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
606 AsnLysGlyArgLysGlyProLeuSerSerIleLeuArgAlaPhe 620

RESULT 13
Q94LX2_SOYBN
ID Q94LX2_SOYBN PRELIMINARY; PRT; 605 AA.
AC Q94LX2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Beta-conglycinin alpha subunit.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
OC Glycine.
OC NCBI_TaxID=3847;
RN [1]_TaxID=3847;
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21327318; PubMed=11434464; DOI=10.1266/ggs.76.99;
RA Yoshino M., Kanazawa A., Teutsuni K., Nakamura I., Shimamoto Y.;
RT "Structure and characterization of the gene encoding alpha subunit of
RT soybean beta-conglycinin.";
RL Genes Genet. Syst. 76:99-105(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=1731988;
RA Lelievre J.M., Dickinson C.D., Dickinson L.A., Nielsen N.C.;
RT "Synthesis and assembly of soybean beta-conglycinin in vitro.";
RL Plant Mol. Biol. 18:259-274(1992).
DR EMBL; AR051865; BAB56161.1; -; Genomic_DNA.
DR PIR; S20007; S20007.
DR HSSP; P25974; 11PJ.
DR SMR; Q94LX2; 194-581.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin_region.
DR Pfam; PF00190; Cupin_1; 2.
SQ SEQUENCE 605 AA; 70306 MW; 8ACB6F8532662984 CRC64;

Alignment Scores:
Pred. No.: 4,49e-76 Length: 605
Score: 1282.00 Matches: 285
Percent Similarity: 59.4% Conservative: 108
Best Local Similarity: 43.1% Mismatches: 173
Query Match: 35.8% Indels: 96
DB: 2 Gaps: 18

US-10-728-323-1 (1-2032) x Q94LX2_SOYBN (1-605)
QY 50 ATGAGAGGAGGAGGTTTCTCCACTGATGCTGTCTAGGATTCCTTCTCGGCTTCAGTT 109
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2 MetArgAlaArgPhePro-----LeuLeuLeuGlyLeuValPheLeuAlaSerVal 19
QY 110 TCTGCAACGCATCCCAAGTCATCACCTTACCAGAAAGAACAGAGAACCCCTCGGCCAG 169
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
20 SerValSerPheGlyIleAla-----TyrTrpGluLys---GluAsnProLysHisAsn 36
QY 170 AGGTGCTCCAGAGTTGTCAACAGGAACCCGATGACTTGAAGCAAAAGGCATCGAGTCT 229
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

Db 37 LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 56
QY 230 CGCTGCACCAAGCTCCAGTATGAT-----CCTCGTTGT 262
Db 57 ArgCysAsnLeuLeuLysValGluLysGluGluCysGluGluGluLeuProArg--- 75
QY 263 GTCTATGATCTCGAGGACACACTGGCACCAACCAACGTTCCCTCAGGGGAGCGG 322
Db 76 -----ProArgProArg-----ProGlnHisProGlu 84
QY 323 ACACGTGGCGCCCAACCCGAGACTACGATGATGAC----- 358
Db 85 ArgGluProGlnGlnProGlyGluLysGluGluAspGluAspGluProArgProIle 104
QY 359 -----CGCGCTCAACCCCGAGAGAG-----GAAGGAGCGCGA 391
Db 105 ProPheProArgProGlnProArgGlnGluGluHisGluGluArgGluGlnGlu 124
QY 392 TGG-----GGACCAAGCTGGACCGAGGAGCGGTGAAGAGAGAA 430
Db 125 TrpProArgLysGluGluLysArgGlyGluLysGlySerGluGluGluAspGlu 144
QY 431 GACTGGAGACAACCAAGAGAGAT-----TGGAGGCGGACCAAGTCAACGACGACCG 484
Db 145 AspGluGlnAspGluArgGlnPheProPheProArgProGlnHisGlnLysGluGlu 164
QY 485 AAAATAAGGCCCGAAG 544
Db 165 ArgLysGlnGluAspGluAspGluGlnGlnArgGluSerGluGluSerGluGluAsp 184
QY 545 AGGGAA-----GAAACATCTCGGAAACACCTTTCTACTTCCCGTCAAGGCGTTAGC 598
Db 185 SerGluLeuArgArgHisLysAsnLysAsnProPheLeuPheGlySerAsnArgPheGlu 204
QY 599 ACCCGTACGGNACCAAAACGCTAGGATCCGGTCTCGAGAGGTTGACCAAGAGTCA 658
Db 205 ThrLeuPheLysAsnGlnTyrGlyArgIleArgValLeuGlnArgPheAsnGlnArgSer 224
QY 659 AGGCAGTTTCAGAAATCTCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTAACCT 718
Db 225 ProGlnLeuGlnAsnLeuArgAspTyrArgIleLeuGluPheAsnSerLysProAsnThr 244
QY 719 CTTGTTCTCCCAAGCAGCTGATCTGATAACATCTTGTATTCCAGCAAGGCGAAGCC 778
Db 245 LeuLeuLeuProAsnHisAlaAspAlaAspTyrLeuIleValIleLeuAsnGlyThrAla 264
QY 779 ACCGTACCGTAGCAATGCGCAATACAGAAAGCTTTAATCTTCAGGAGGCCATGCA 838
Db 265 IleLeuSerLeuValAsnAsnAspAspArgAspSerTyrArgLeuGlnSerGlyAspAla 284
QY 839 CTGAGATCCCATCCCGTTTCATTCTCTACATCTTCAACCGCCATGACCAACGAGACCTC 898
Db 285 LeuArgValProSerGlyThrThrTyrTyrValValAsnProAspAsnAsnGluAsnLeu 304
QY 899 AGAGTAGTAAATCTCCATCCCGTTAAACACACCCCGCAGTTTGAGGATTTCTTCCCG 958
Db 305 ArgLeuIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSerPhePheLeu 324
QY 959 GCGAGCAGCGGACCAATCTCTACTTCAGCGGCTTCAGCGAGTAATAGTTGGAGGCC 1018
Db 325 SerSerThrGluAlaGlnSerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAla 344
QY 1019 GCCTTCAATCGGAAATTCATAGATACGAGGAGTCTGTAGAGAGAATGCGAGAGT 1078
Db 345 SerTyrAspThrLysPheGluGluIleAsnLysValLeuPheSerArgGluGluGly--- 363
QY 1079 GAGCAAGAGGAGAGCGGAGCGGCTGAGTACTCGGAGTAGTGAGAACATGAAGA 1138
Db 364 -----GlnGlnGlnGlyGluGlnArgLeu-----GlnGluSer 374
QY 1139 GTGATAGTCAAGTGTCAAGGAGCAGCTGTGAAGAACTTACTTAAGCAGCCTAAATCCGTC 1198
Db 375 ValIleValIleSerLysGluGlnIleArgAlaLeuSerLysArgAlaLysSerSer 394

QY 1199 TCAAGAAGAGGCTCCGAAGAGAGAGGAGATATCACCACCCCAATCAACTTGAGAGAGC 1258
Db 395 SerArgLysThrIleSerSerGlu-----AspLysProPheAsnLeuArgSerArg 411
QY 1259 GAGCCCGATCTTTCTAAACAACCTTTGGGAAGTTATTGAGGTGAAGCCACAGAGAAGAC 1318
Db 412 AspProIleTyrSerAsnLysLeuGlyLysPhePheGluIleThrProGlu---LysAsn 430
QY 1319 CCCAGCTTCAGACCTGGACATGATCTCCTGTGTAGAGATCAAAAGAAGAGAGCTTTG 1378
Db 431 ProGlnLeuArgAspLeuAspIlePheLeuSerIleValAspMetAsnGluGlyAlaLeu 450
QY 1379 ATGCTCCCACTTCAACTCAAAAGCCATGTTTATCGTCTCGTCAACAAAGAACTGGA 1438
Db 451 LeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGlyAspAla 470
QY 1439 AACCTTGAACCTCGTGGCTGTAAAGAAAGACCAACAGAGGGGACGGGGAGAGAG 1498
Db 471 AsnIleGluLeuValGlyLeu---LysGluGlnGlnGlnGlnGlnGlnGluGln 489
QY 1499 GAGGACGAGACGAAGAGAGGAGGAGTAACAGAGAGGTGCTAGGTACACAGCGAGG 1558
Db 490 Pro-----LeuGluValArgLysTyrArgAlaGlu 499
QY 1559 TTGAAGGAGGCGATGTGTTTCATCATCCAGCAGCTCATCCAGTAGCCATCAACGTTCC 1618
Db 500 LeuSerGluGlnAspIlePheValIleProAlaGlyTyrProValValAsnAlaThr 519
QY 1619 TCGGAATCTCATCTCGTTCGGTTCATCAACGCTGAAAGAAACACACAGAAATCTTCTT 1678
Db 520 SerAsnLeuAsnPhePheAlaIleGlyIleAsnAlaGluAsnAsnGlnArgAsnPheLeu 539
QY 1679 CGAGGTGATAAGGACAATGTGATAGACAGATAGAGAACCAAGCAGAGATTAGCATTC 1738
Db 540 AlaGlySerGlnAspAsnValIleSerGlnIleProSerGlnValGlnGluLeuAlaPhe 559
QY 1739 CTTGGTTCGGGTGAACAAAGTTGAGAGGCTCATCAAAACCAAGAGAAATCTCACTTTGTG 1798
Db 560 ProGlySerAlaGlnAlaValGluLysLeuLeuLysAsnGlnArgGluSerTyrPheVal 579
QY 1799 AGTGTCTGCTCAATCTCAATCTCAATCTCAATCTCGTCTCTCGAGAGAGTCTCTCGAG 1858
Db 580 AspAlaGlnProLysLys----- 585
QY 1859 AAAGAGGATCAAGAGGAGGAGAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1918
Db 586 -----LysGluGluGlyAsnLysGlyArgLysGlyProLeuSerSerIleLeuArg 602
QY 1919 GCTTTT 1924
Db 603 AlaPhe 604

RESULT 14

GLCA, SOYBN

ID GLCA SOYBN STANDARD; PRT; 605 AA.

AC P13916;

DT 01-JAN-1990 (Rel. 13, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE Beta-conglycinin, alpha chain precursor.

OS Glycine max (Soybean).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

OC Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

OC Glycine.

OX NCBI_TaxID=3847;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Cotyledon;

RX MEDLINE=91355860; PubMed=2103438;

RA Sebastiani F.L., Farrel L.B., Schuler M.A., Beachy R.N.;

RT "Complete sequence of a CDNA of alpha subunit of soybean beta-

RT conglycinin, ";
 RL Plant Mol. Biol. 15:197-201 (1990).
 CC -I- FUNCTION: Seed storage protein. Accumulates during seed
 CC development and is hydrolyzed after germination to provide a
 CC carbon and nitrogen source for the developing seedling.
 CC -I- SUBUNIT: The alpha', alpha", and beta-subunits associate in
 CC various combinations to form trimeric proteins.
 CC -I- SUBCELLULAR LOCATION: Embryo axis, and cotyledonary membrane-bound
 CC vacuolar protein bodies.
 CC -I- SIMILARITY: Belongs to the 7S seed storage protein family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; X17698; CAA35691.1; -; mRNA.
 DR PIR; S14681; FWSYBA.
 DR HSSP; P25974; IIPJ.
 DR SNR; P13916; 194-381.
 DR InterPro; IPR006045; Cupin.
 DR InterPro; IPR007113; Cupin region.
 DR Pfam; PF00190; Cupin_1; 2.
 KW Glycoprotein; Multigene family; Seed storage protein; Signal;
 KW Storage protein; Vacuole.
 FT SIGNAL 1 22
 FT PROPEP 23 62
 FT CHAIN 63 605 Beta-conglycinin, alpha chain.
 FT CARBOHYD 261 261 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 517 517 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 605 AA; 70293 MW; CBEBAS0506BBBC57 CRC64;
 Alignment Scores:
 Pred. No.: 5,22e-76 Length: 605
 Score: 1281.00 Matches: 285
 Percent Similarity: 59.2% Conservative: 107
 Best Local Similarity: 43.1% Mismatches: 174
 Query Match: 35.7% Indels: 96
 DB: 1 Gaps: 18
 US-10-728-323-1 (1-2032) x GLCA_SOYBN (1-605)
 QY 50 ATGAGAGGAGGAGGTTCTCCACTGATGCTGCTCTAGGATCTTGTCTGGGTTTCAGTT 109
 DB 2 MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyLeuValPheLeuAlaSerVal 19
 QY 110 TCTGCAACGATGCCAAGTCACTACCTTACCAGAAGAAACAGAGAACCCCTCGGCCAG 169
 DB 20 SerValSerPheGlyIleAla-----TyrTrpGluLys-----GluAsnProLysHisAsn 36
 QY 170 AGGTGCTCAGAGTTGTCAACAGGACCGGATGACTTGAAGCAAAAGGATCGGATCT 229
 DB 37 LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 56
 QY 230 CGCTGCACCAAGCTCAGTATGAT-----CTCTGTTGT 262
 DB 57 ArgCysAsnLeuLeuLysValGluLysGluCysGluGluGlyLeuPheProArg--- 75
 QY 263 GTCTATGATCTCGAGGACACTGGCCACCACCAACAGCTTCCCTCCAGGGGAGCGG 322
 DB 76 -----ProArgProArg-----ProGlnHisProGlu 84
 QY 323 ACACGTGGCGCAACCCGAGACTACGATGATGAC----- 358
 DB 85 ArgGluProGlnProGlyGluLysGluLysGluLysGluLysGluLysGluLysGluLys 104
 QY 359 -----CGCGCTCAACCCCAAGAGAG-----GNAGAGGCCGA 391
 DB 105 ProPheProArgProGlnProArgGlnGlnGluHisGluGlnArgGluGlnGlu 124
 QY 392 TGG-----GGACCAAGCTGGACCGAGGGAGCGTGAAGAGAGAA 430
 DB 471 AsnIleGluLeuValGlyLeu---LysGluGlnGlnGlnGlnGlnGlnGlnGlnGln 489

Db 125 TrpProArgLysGluGluLysArgGlyGluLysGlySerGluGluGluAspGluAspGlu 144
 QY 431 GACTGAGAGCAACCAAGAGAAAT-----TGAGCGCAGCAAGTCTATCAGCAGCCACGG 484
 Db 145 AspGluGluAspGluArgGlnPheProPheProArgProHisGlnLysGluGlu 164
 QY 485 AAAATAAGCCCGCAAGGAGAGCAAGAGTGGGGAACACACAGGTAGCCATGTG 544
 Db 165 ArgAsnGluGluGluAspGluGluGlnArgGluSerGluGluSerGluAsp 184
 QY 545 AGGGAA-----GAAACATCTCGGAACACCCCTTCTCTCCGTCGAGCGGTTTAC 598
 Db 185 SerGluLeuArgArgHisLysAsnLysAsnProPheLeuPheGlySerAsnArgPheGlu 204
 QY 599 ACCCGTACGGGAACCAAAACCGTAGGATCCGGTCTCGAGAGTTTCCACCAAGGTCA 658
 Db 205 ThrLeuPheLysAsnGlnTyrGlyArgIleArgValLeuGlnArgPheAsnGlnArgSer 224
 QY 659 AGGCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTAACACT 718
 Db 225 ProGlnLeuGlnAsnLeuArgAspTyrArgIleLeuGluPheAsnSerLysProAsnThr 244
 QY 719 CTTGTTCTTCCCAAGCACGCTGATGCTGATAACATCTTTGTTATCCAGCAAGGGCAAGCC 778
 Db 245 LeuLeuLeuProAsnHisAlaAspAlaAspTyrLeuIleValIleLeuAsnGlyThrAla 264
 QY 779 ACCGTGACCGTAGCAATCGCAATACAGAAAGAGCTTTAATCTTCACGAGGGCCATGCA 838
 Db 265 IleLeuSerLeuValAsnAsnAspArgAspSerTyrArgLeuGlnSerGlyAspAla 284
 QY 839 CTCAGAATCCCATCCGGTTTCATTCCTACATCTTGAAACCGCATCACACCAAGAACCTC 898
 Db 285 LeuArgValProSerGlyThrTyrTyrValValAsnProAsnAsnGlnAsnLeu 304
 QY 899 AGAGTAGCTAAAATCTCCATGCCGCTTAAACACACCCGCGCAGTTTCAGGATTTCTCCCG 958
 Db 305 ArgLeuIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSerPhePheLeu 324
 QY 959 GCGAGCAGCGGAGACCAATCATCTTCTGAGGGGTTTCAGAGGAATAGTTGGAGGCC 1018
 Db 325 SerSerThrGluAlaGlnGlnSerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAla 344
 QY 1019 GCCTTCAATCGGAATTCATGAGATACGAGGCTCTCTTAGACAGAAATCGAGAGGT 1078
 Db 345 SerTyrAspThrLysPheGluIleAsnLysValLeuPheSerArgGluGly--- 363
 QY 1079 GAGCAAGAGAGAGAGCGGAGCGGATCGAGTACTCGGAGTAGTGAGAACATGAAGGA 1138
 Db 364 -----GlnGlnGlnGlyGluGlnArgLeu-----GlnGluSer 374
 QY 1139 GTGATAGTCAAAGTCAAAGGAGCAGCTTGAAGAACTTACTTAAGCAGCTAAATCCGTC 1198
 Db 375 ValIleValIleSerLysGluGlnIleArgAlaLeuSerLysArgAlaLysSerSer 394
 QY 1199 TCAAGAAAAGGCTCCCAAGAGAGGGAGATATCACCAACCCCAATCACTTGAGAGAGGC 1258
 Db 395 SerArgLysThrIleSerSerGlu-----AspLysProPheAsnLeuArgSerArg 411
 QY 1259 GAGCCCGATCTTTCTAAACAACCTTTGGGAAAGTTATTGAGGTGAAGCAGACAAAGAAG 1318
 Db 412 AspProIleTyrSerAsnLysLeuGlyLysPhePheGluIleThrProGlu---LysAsn 430
 QY 1319 CCCAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAGAGAGCTTTG 1378
 Db 431 ProGlnLeuArgAspLeuAspIlePheLeuSerIleValAspMetAsnGluGlyAlaLeu 450
 QY 1379 ATGCTCCCACTTCAACTCAAAGGCGCATGTTTCGTCTGCTCAACAAAGAACTGGA 1438
 Db 451 LeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGlyAspAla 470
 QY 1439 AACCTTGAATCGTGTGCTTAAGAAAAGACAAACAGAGGGGAGCGCGGGAAGAGAG 1498
 Db 471 AsnIleGluLeuValGlyLeu---LysGluGlnGlnGlnGlnGlnGlnGlnGlnGln 489


```

Db      226 ThrValLeuSerProAsnAspArgAsnSerTyrAsnLeuGluArgGlyAspThrIleLys 245
      ||||| ::: |||::||| |||::|||::|||::|||::||| :||| :||| :|||::
QY      845 ATCCCATCCGGTTTCATTTCTCATCTTGAACCGCCATGCACACAGAACCTCAGAGTA 904
      |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      246 IleProAlaGlyThrSerTyrLeuValAsnGlnAspAspGluGluAspLeuArgVal 265
      |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      905 GCTAAATCTCCATGCCGTTAAACACACCCGGCCAGTTTGAGATTTCCTCCCGGCGAGC 964
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      266 ValAspPheValIleProValAsnArgProGlyLysPheGluAlaPhe-----GlyLeu 283
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      965 AGCCGAGACCAATCATCTCTTGCAGCGCTTCAGCAGGAATACGTTGGAGGCGCGCTTC 1024
      |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      284 SerGluAsnLysAsnGlnIleGluArgGlyPheSerIysAsnIleLeuGluAlaSerLeu 303
      |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      1025 AATGCGGAATTCATGAGATACGAGGGTCTCTTAGAAGAGAAATCGAGAGGTGAGCAA 1084
      |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      304 AsnThrLysTyrGluThrIleGluLysValLeuLeuGluGlnGluLysLysProGln 323
      |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      1085 GAGGAGAGGCGAGGCGATGGAGTACTCGGAGTAGTGAGAACATGAAGAGTGATA 1144
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      324 GlnLeuArgAspArgLysArg-----ThrGlnGlnGlyGluGluArgAsp---AlaIle 340
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      1145 GTCAAGTGTCAAGGAGCAGCTTGAAGAACTTACTAAGCACGCTAAATCCGCTCAAAG 1204
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      341 IleLysValSerArgGluGlnIleGluLysValLeuLysLysSerSerLys 360
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      1205 AAAGGCTCCGAAGAAGGAGGATATCACCAACCCCAATCACTTGAGAGAGCGAGCCC 1264
      |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      361 LysSerLeuProSerGluPheGlu-----ProPheAsnLeuArgSerHisLysPro 377
      |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      1265 GATCTTTCTAAACAATTTGGGAAGTTATTGAGGTGAAGCCAGACACAAGAAGACCCCGAG 1324
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      378 GluTyrSerAsnLysPheGlyLysLeuPheGluIleThrProGluLysLysTyrProGln 397
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      1325 CTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAAGGAGCTTTGATGCTC 1384
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      398 LeuGlnAspLeuAspIleLeuValSerCysValGluIleAsnLysGlyAlaLeuMetLeu 417
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      1385 CCACATTTCAACTCAAGGCCATGGTTATCGTCGTCTCAACAAGGAACTGGAAACCTT 1444
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      418 ProHisTyrAsnSerArgAlaIleValLeuLeuValAsnGluGlyLysGlyAsnLeu 437
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      1445 GAATCTCGTCTGTAAAGAAAGCAACAACAGAGGGGACGGCGGGAAGAGGAGGAC 1504
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      438 GluLeuLeuGlyLeuLysAsnGluGlnGlnArg-----449
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      1505 GAAGACGAAGAAGAGGGAAGTAACAGAGAGGTGCGTAGGTACACAGCGAGGTTGAAG 1564
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      450 GluAspArgLysGluArg-----AsnAsnGluValGlnArgTyrGluAlaArgLeuSer 467
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      1565 GAAGGCGATGTGTTTCATGCGCAGCAGCTCATCCAGTAGCCATCAACGTTCTCCCGAA 1624
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      468 ProGlyAspValValIleIleProAlaGlyHisProValAlaIleSerAlaSerSerAsn 487
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      1625 CTCATCTGTTGCTTTCGGTTCACGCTCAAAACAACCAACACAGAACTCTTCCTGCAGGT 1684
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      488 LeuAsnLeuLeuGlyPheGlyIleAsnAlaLysAsnAsnGlnArgAsnPheLeuSerGly 507
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      1685 GATAAGGACAAATGTGATAGACCATAGAGAAGCAAGCAAGGAGATTTCCTCTGGG 1744
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      508 SerAspAsnValIleSerGlnIleGluAsnProValLysGluLeuThrPheProGly 527
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      1745 TCGGGTGAACAAGTTGAGAGCTCATCAAAAACAGAGGAATCTCATCTTTGTGAGTGCT 1804
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      528 SerSerGlnGluValAsnArgLeuIleLysAsnGlnLysGlnSerHisPheAlaSerAla 547
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      1805 CGTCTCAATCTCAATCTCAATCTCCGTCGTCCTCGAAGAGAGTCTCCTGAGAAAGAG 1864
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db      548 GluPro-----Glu 550
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      1865 GATCAAGAGGAGGAGAAACCAAGGAGGAGGTCCTCTTTCATTTTGAAGGCTTTT 1924
      :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
```

Db 551 GlnLysGluGluSerGlnArgLysArgSerProLeuSerSerValLeuAspSerPhe 570

Search completed: May 15, 2006, 22:18:16
Job time : 502.942 secs

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 23:34:07 ; Search time 57.3506 Seconds
(without alignments)
4441.255 Million cell updates/sec

Title: US-10-728-323-1

Perfect score: 3586

Sequence: 1 aataatcatatattcatc.....cgttgtgctgtttcttc 2032

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p model -DRV=xlp
-Q=/abs/ABSWEB_spool/US10728323/runat_15052006_172145_22607/app_query.fasta_1
-DB=Published Applications AA_Main -Qfmt=fastan -SUFFIX=n2p.rapbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=biosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US10728323 @CGN_1_1_503 @runat_15052006_172145_22607 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPELCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA_Main.*

- 1: /cgn2_6/prodata/1/pubpaa/us07_PUBCOMB.pcp.*
- 2: /cgn2_6/prodata/1/pubpaa/us08_PUBCOMB.pcp.*
- 3: /cgn2_6/prodata/1/pubpaa/us09_PUBCOMB.pcp.*
- 4: /cgn2_6/prodata/1/pubpaa/us10A_PUBCOMB.pcp.*
- 5: /cgn2_6/prodata/1/pubpaa/us10B_PUBCOMB.pcp.*
- 6: /cgn2_6/prodata/1/pubpaa/us11_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3286	91.6	626	3	US-09-847-208-28
2	3286	91.6	626	4	US-10-228-806-2
3	3286	91.6	626	4	US-10-100-303A-7
4	3286	91.6	626	4	US-10-245-871-10
5	3286	91.6	626	4	US-10-253-286-10
6	3286	91.6	626	5	US-10-809-689-95
7	3286	91.6	626	5	US-10-899-551-2
8	3192	89.0	635	5	US-10-899-551-53
9	3052	85.1	634	3	US-09-731-221-78
10	3041	84.8	614	3	US-09-331-631A-21
11	3041	84.8	614	3	US-09-847-208-27

12	3041	84.8	614	4	US-10-147-095-21	Sequence 21, Appl
13	3041	84.8	614	4	US-10-100-303A-8	Sequence 8, Appl
14	1394	38.9	268	4	US-10-100-303A-55	Sequence 55, Appl
15	1282	35.8	605	4	US-10-424-599-153195	Sequence 153195
16	1281	35.7	605	3	US-09-331-631A-25	Sequence 25, Appl
17	1281	35.7	605	4	US-10-147-095-25	Sequence 110, Appl
18	1275.5	35.6	623	5	US-10-424-599-153206	Sequence 153206
19	1274.5	35.5	623	5	US-10-739-930-8840	Sequence 8840, Ap
20	1268	35.4	605	4	US-10-100-303A-110	Sequence 110, App
21	1201.5	33.5	584	4	US-10-424-599-260105	Sequence 260105
22	1157	32.3	417	4	US-10-245-227B-14	Sequence 14, Appl
23	1157	32.3	425	4	US-10-245-227B-12	Sequence 12, Appl
24	1156	32.2	439	4	US-10-245-227B-1	Sequence 1, Appl
25	1117.5	31.2	390	4	US-10-245-227B-2	Sequence 2, Appl
26	874	24.4	324	4	US-10-425-114-44408	Sequence 44408, A
27	867	24.2	344	4	US-10-425-114-51703	Sequence 51703, A
28	865.5	24.1	666	3	US-09-331-631A-3	Sequence 3, Appl
29	865.5	24.1	666	4	US-10-147-095-3	Sequence 3, Appl
30	863.5	24.1	625	3	US-09-331-631A-5	Sequence 5, Appl
31	863.5	24.1	625	4	US-10-147-095-5	Sequence 5, Appl
32	849.5	23.7	666	3	US-09-331-631A-1	Sequence 1, Appl
33	849.5	23.7	666	4	US-10-147-095-1	Sequence 1, Appl
34	828	23.1	316	4	US-10-425-114-43905	Sequence 43905, A
35	828	23.1	316	4	US-10-425-114-49262	Sequence 49262, A
36	827	23.1	315	4	US-10-425-114-43971	Sequence 43971, A
37	816	22.8	301	4	US-10-425-114-49373	Sequence 49373, A
38	791	22.1	590	3	US-09-331-631A-8	Sequence 8, Appl
39	791	22.1	590	4	US-10-147-095-8	Sequence 8, Appl
40	790	22.0	300	4	US-10-425-114-43597	Sequence 43597, A
41	788	22.0	296	4	US-10-425-114-43984	Sequence 43984, A
42	770	21.5	291	4	US-10-425-114-51411	Sequence 51411, A
43	761	21.2	149	4	US-10-100-303A-54	Sequence 54, Appl
44	737.5	20.6	525	3	US-09-331-631A-7	Sequence 7, Appl
45	737.5	20.6	525	4	US-10-147-095-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1

US-09-847-208-28
; Sequence 28, Application US/09847208
; Publication NO. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; FILE REFERENCE: Ige-MEDIATED ALLERGIC DISEASES
; FILE REFERENCE: UC67.002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Arachis hypogaea (Peanut)
US-09-847-208-28

Alignment Scores:
Pred. No.: 2.56e-261 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.6% Indels: 0
DB: 3 Gaps: 0

US-10-728-323-1 (1-2032) x US-09-847-208-28 (1-626)

QY 50 ATGAGAGGAGGTTTCTCCACTGATGCTGTGCTAGGATCCTTCTGCTTCAGTT 109

Db 1 MetArgGlyArgValSerProLeuLeuLeuGlyLeuValLeuAlaSerVal 20

QY 110 TCTGCAACGATGCCAAGTCATCACTTACCAGAGAAACACAGAGAACCCCTCGCCCGCAG 169
Db 21 SerAlaThrHisAlaLysSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGGTGCTCCAGAGTTGTCAACAGGACCGGATGACTTGAAGCAAAAGGCATCGAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGTGCACCAAGCTCGAGTATGATCCTCGTTGTCTATGATCCTCGAGACACACTGGC 289
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAAGTTCCTCCAGGGGACCGACACAGTGGCCGCCCAACCCGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCTCAACCCCAAGAGAGGAAGGAGCGCGATGGGACCCAGCTGGACCG 409
Db 101 AspAspArgArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro 120
QY 410 AGGAGCGTGAAGAGAAGACTGGAGACAAACCAAGAGAAGATTGGAGCGACCAAGT 469
Db 121 ArgGluArgGluArgGluArgTyrArgGlnProArgGluAspTyrArgA:gpProSer 140
QY 470 CATCAGCAGCCAGGAAATAAGCCCGAAGGAGAGAGAGAGAACAGAGTGGGAACA 529
Db 141 HisGlnProArgLysIleArgProGluGlyArgGluGluGlnGluTyrGlyThr 160
QY 530 CCAAGTAGCCATGTAGGGAAGAAACATCTCGGAACAAACCTTCTACTTCCCGTCAAGG 589
Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCCGCTACGGGAACCAAAACGGTAGGATCCGGTCTCGCAGAGGTTTGAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGTCAAGGAGTTTCAGATCTCCAGATCACCCTATGTCAGATCGAGGCCAAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAAACACTCTGTCTTCCCAAGCAGCTGATGCTGATAACATCTTGTATCCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTACCGTAGCAAAATGGCAATAACAGAAAGACTTAACTTTCAGCAG 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCATGCATCAGATCCCATCCGTTTTCATTTCTTACATCTTGACCGCCCATGACAAC 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCCCGTTAAACACACCCGCCAGTTTGAGGAT 949
Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGGAGCAGCAGCAACATCATCTTACTTTCAGGCTTTCAGCAGAAATACG 1009
Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCCGCTTCATGCGGAATTCATAGATAGCGAGGCTGCTGTAGAGAGAAT 1069
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGGAGAGGCGCAGGCGATGGAGTACTCGGAGTAGTGAAGAC 1129
Db 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGAGTATAGTCAAGTGTCAAAGTGTCAAAGAGACCGTTCGAAGAACTTACTAAGACGCT 1189
Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCCGTCTCAAGAAAGGCTCCGAAGAGAGGAGATATCACCACCCCAATCAACTTG 1249

Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAGGCGAGCCCGCATCTTCTTAACAACATTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAGAACCCCGAGCTTCAGACCTTGACCATGTCTCACCTGTGTAGAGATCAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAAGGCCATGTTTATCTCTCGTCTCGTCAACAAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
QY 1430 GGAACCTGGAACCTTGAACCTCGTGGCTGTGAAGAAAAGACAAACACAGAGGGGACCGCG 1489
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyA:sgArg 480
QY 1490 GAAGAAGAGAGACGAAGACGAAGAGAGGAGGAGTAACACAGAGAGGTGCGTAGGTAC 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTGAAGGAAGCGCATGTTCATCATCCACAGCAGCTCATCCAGTAGGCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAACCTCCATCTGCTTGGCTTCGCTATCAACGCTCAAAACACACACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTCGAGGTGATTAAGACAAATGTATAGACCATAGACAGAGGAGCAAGGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTGCGGTGAACAAAGTTGAGAAGCTCATCAAAACACAGAGGAATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTCTCGTCTCTCAATCTCAATCTCAATCTCCGCTGCTCTCTGAGAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTCAGAAAGAGATCAAGAGGAGAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
Db 601 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTTAAC 1927
Db 621 IleLeuLysAlaPheAsn 626

RESULT 2

US-10-228-806-2
; Sequence 2, Application US/10228806
; Publication No. US20030049237A1
; GENERAL INFORMATION:
; APPLICANT: Bannon, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reactions
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0043
; CURRENT APPLICATION NUMBER: US/10/228, 806
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-228-806-2

Alignment Scores: 2.56e-261 Length: 626
Pred. No.: 3286.00 Matches: 626
Score:

Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.6% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-1 (1-2032) x US-10-228-806-2 (1-626)

QY 50 ATGAGAGGAGGTTTCTCCACTGATGCTGTTCTAGGATCCTTCTCGGCTTCAGTT 109
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyLeuValLeuAlaSerVal 20

QY 110 TCTGCAACGATCCCAAGTCATCACCTTACCAGAGAAACAGAGAACCCCTGCGCCAG 169
DB 21 SerAlaThrHisAlaLysSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40

QY 170 AGGTGCTCCAGAGTTGTCAACAGAACCGGATGCTTGAAGCAAAAGGCATGCGAGTCT 229
DB 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60

QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTTGTCTATGATCCTCGAGGACACACTGGC 289
DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80

QY 290 ACCACCAACCAAGTTCCTCCAGGGAGCGGACAGTGGCGGCCCAACCCGAGACTAC 349
DB 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100

QY 350 GATGATGACCGCTCAACCCCGAAGAGAGGAGGCGGATGGGGACCGTGGACCG 409
DB 101 AspAspArgArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro 120

QY 410 AGGGAGCGTGAAGAGAGAAGACTGGAGACCAACAGAGAGATTTGGGGGACCAAGT 469
DB 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgProSer 140

QY 470 CATCAGCAGCCGGAATAAGGCCGGAAGAGAGAGAGAGAACAGAGTGGGAAACA 529
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160

QY 530 CAGGTAGCCTGTGAGGAGAGAAACATCTCGGAACAACCTTCTACTTCCCGTCAAGG 589
DB 161 ProGlySerHisValArgGluGluThrSerArgAsnProPheTyrPheProSerArg 180

QY 590 CGGTTTACGACCGCTACGGGAACCAAAACGGTAGGATCCGGTCTCGAGAGGTTTGAC 649
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200

QY 650 CAAAGTCAAGCAGTTTCAGAAATCTCAGAAATCACCGTATTGTGAGATCGAGGCCAAA 709
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220

QY 710 CCTAACACTCTTCTTCCAGACGCTGATGCTGATAACATCTTCTTATCCAGCAA 769
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleuValIleGlnGln 240

QY 770 GGGCAAGCCACCGTAGCCGTAGCAATGGCAATAACAGAAAGAGCTTTAATCTTCACGAG 829
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260

QY 830 GGGCATGCACTCAGATCCATCCGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 889
DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280

QY 890 CAGAACTCAGAGTACGTAATCTCCATGCCGTTTAAACACACCCCGCCAGTTTGGAGAT 949
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300

QY 950 TTCTTCCCGGAGCAGCCGAGACCAATCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1009
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320

QY 1010 TTGAGGCGCCCTTCAATCGGAAATTCATATGATACGAGGCTGTGTAGAGAAAT 1069

DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn 340

QY 1070 GCAGGAGGTGAGCAAG 1129

DB 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgArgArgArgArgArgArgArgArgArg 360

QY 1130 AATGAAGGAGTATAGTCAAAAGTGTCAAGAGAGACGTTTGAAGAACTTACTTAAGCACGCT 1189

DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380

QY 1190 AATCCGCTCAAGAAAGAGTCCCAAGAGAGAGAGAGATATCACCAACCAATCAACTTG 1249

DB 381 LysSerValSerLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400

QY 1250 ACAGAAAGGCGAGCCCGATCTTCTTCAACAACTTTGGAACTTATTTGAGGTGAAGCCAGAC 1309

DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420

QY 1310 AAGAAGAAACCCCGCTTACGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369

DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440

QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGAGGCGATGTTATCGTCTCGTCAACAA 1429

DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460

QY 1430 GGAACTGGAAACCTTGAACCTCGTGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1489

DB 461 GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnGlnArgGlyArgArg 480

QY 1490 GAAGAGAGGAGGAT 1549

DB 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500

QY 1550 ACAGCGAGGTTCAGAGAGGCGATGTTTCATCATGCCAGAGCTCATCCAGTAGCCATC 1609

DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520

QY 1610 AACGCTTCTCCGAACTCCATCTGCTTGGCTTTCGCTTCAACGCTGAAACCAACACAGA 1669

DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg 540

QY 1670 ATCTTCTTTCAGGAGTAAAGGACAAATGTGATAGACAGATAGAGAGAGAGAGAGAGAT 1729

DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560

QY 1730 TTAGCATTCCTCGGTCGGTGAAACAAGTTGAGAGAGCTCATCAAAACCCAGAGAGATCT 1789

DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580

QY 1790 CACTTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCT 1849

DB 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGlyLysGlu 600

QY 1850 TCTCTGAGAAAGAGATCAAGAGAGGAGAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAT 1909

DB 601 SerProGluLysGluAspGlnGluGluGluGlnGlyLysGlyProLeuLeuSer 620

QY 1910 ATTTTGAAGGCTTTTAAAC 1927

DB 621 IleLeuLysAlaPheAsn 626

RESULT 3

US-10-100-303A-7
; Sequence 7, Application US/10100303A
; Publication No. US20030202980A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100,303A
; CURRENT FILING DATE: 2002-03-18

; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Atachis hypogaea, Prot/Nucleo-Ara h 1
US-10-100-303A-7

Alignment Scores:
Pred. No.: 2,566-261 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.6% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-1 (1-2032) x US-10-100-303A-7 (1-626)

QY	50	ATGAGAGGGAGGCTTTCTCCACTGATGCTGTTCTAGGGATCCTTCTCTGGCTTCAGTT	109
DB	1	MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal	20
QY	110	TCTGCAACGATGCCAAGTCATCACCTTACCAGAAGAAACACAGAACCCCTCGCCACG	169
DB	21	SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln	40
QY	170	AGTGTCTCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCGATCGAGTCT	229
DB	41	ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer	60
QY	230	CGTGTCCACCAAGCTCCAGTATGATCTCGTTGTCTATGATCCTCGAGGACACACTGGC	289
DB	61	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly	80
QY	290	ACCACCAACCAACGTTTCCCTCCAGGGAGCGGACACGTCGGCCCAACCCGAGACTAC	349
DB	81	ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	100
QY	350	GATGATGACCGCGTCAACCCGGAAGAGAAAGGAGCGCGATGGGACACAGCTGACCG	409
DB	101	AspAspArgArgGlnProArgGluGluGlyArgTyrGlyArgTyrProAlaGlyPro	120
QY	410	AGGAGCGTCAAGAGAGAGAGACTGGAGACACCAAGAGAGATTTGGAGCGGACCAAGT	469
DB	121	ArgGluArgGluArgGluGluAspTyrArgGlnProArgGluAspTyrArgProSer	140
QY	470	CATCAGCAGCCCGGAAAATAAGCCCGAAGGAGAGAGAGAACAAAGAGTGGGAACA	529
DB	141	HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGlnGluTyrGlyThr	160
QY	530	CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAAACCCCTTCTACTTCCCGTCAAGG	589
DB	161	ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg	180
QY	590	CGGTTTAGCACCCGTACGGGAACCAAAACGGTAGGATCCGGTCTCTGACAGAGTTTGAC	649
DB	181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	200
QY	650	CAAAAGGTCAAGCGATTTCCAGATCTCCAGAATCACCGTATTGTGACAGATCGAGGCCAA	709
DB	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	220
QY	710	CCTAACACTCTTCTTCCAGCAGCGGTGATGCTGATAACATCTCTTGTATTCACGAA	769
DB	221	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	240
QY	770	GGGCAAGCCACCGTAGCCGATGGCAATGGCAATAACAGAAAGAGCTTTAATCTTTGACG	829
DB	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260
QY	830	GGCCATGCACTCAGAAATCCCATCCGTTTCATTTCTTACATCTTGAACCCCATGACAAC	889
DB	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280

QY	890	CAGAACTCAGAGTAGCTAAATCTCCATGCCGTTAAACACACCCGCGCAGTTTGAGGAT	949
DB	281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsn	300
QY	950	TTCTTCCCGCGAGCAGCGAGACCAATCATCTTCTGACGGCTTTCAGCAGGAAATACG	1009
DB	301	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	320
QY	1010	TTGAGGCGCCTTCAATCGGAAATTCATAGATACGAGGGTGTCTGTAGAGAGAAAT	1069
DB	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn	340
QY	1070	GCAGAGGTGACAAAGAGAGAGAGCGGACGAGCGATCGAGTACGCGAGTAGTGAGAAC	1129
DB	341	AlaGlyGlyGlnGlnGluArgGlyGlnArgArgTyrPsrThrArgSerSerGluAsn	360
QY	1130	AATGAAGGAGTGTATAGTCAAAGTGTCAAAGGAGCACGTTGAAGAACTTAAAGCACGCT	1189
DB	361	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380
QY	1190	AAATCCGTCTCAAGAAAGGCTCCGAAAGAGGAGATATCACCAACCCCAATCAACTTG	1249
DB	381	LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu	400
QY	1250	AGAGAGGCGAGCCGATCTTCTAACACTTTGGGAAGTATTATTCAGGTGAAGCCAGAC	1309
DB	401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420
QY	1310	AAGAAGAACCCCGACTTCAGGACCTGGACATGATGTCTCACCTGTGTAGAGATCAAGAA	1369
DB	421	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu	440
QY	1370	GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGGTTATCGTCTGTCAACAAA	1429
DB	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	460
QY	1430	GGAACTGGAAACTTGAACCTCGTGGCTGTAAAGAAAGAGCAACAACAGAGGGGACGCGG	1489
DB	461	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnArgGlyArgArg	480
QY	1490	GAAGAAGAGGAGCAAGACGAAAGAGGAGGGAAGTAACAGAGAGGTGCGTAGGTAC	1549
DB	481	GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr	500
QY	1550	ACAGCGAGTTGAAGAAAGCGATGTTCATCATGCCAGCAGCTCATCCAGTAGCCATC	1609
DB	501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520
QY	1610	AACGCTTCTCCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTCAACCAACCAACAG	1669
DB	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg	540
QY	1670	ATCTTCTTCCAGGATGATAAGGACAAATGTATAGACCAATAGACAGAGCAAGCAAGGAT	1729
DB	541	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp	560
QY	1730	TTAGCAATTCCTGGGTGCGGTGAACAAAGTTGAGAAGCTCATCAAAAACAGAGGAACTCT	1789
DB	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	580
QY	1790	CACTTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCCGCTCTCTCTGAGAAGAG	1849
DB	581	HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu	600
QY	1850	TCTCTTGAGAAAGAGATCAAGAGGAGGAGAAACCAAGGAGGAGGAGGCTCCATCTCTTCA	1909
DB	601	SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyGlyGlyProLeuLeuSer	620
QY	1910	ATTTCGAAGGCTTTTAAAC	1927
DB	621	IleLeuLysAlaPheAsn	626

Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCGTCTCGTCTCGAAGAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTGTAGAAAGAGATCAAGAGGAGGAAACCAAGGAGGGAAGGTCACCTCTTTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGCTTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626

RESULT 5
US-10-253-286-10
; Sequence 10, Application US/10253286
; Publication No. US20040058881A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: II-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2015
; CURRENT APPLICATION NUMBER: US/10/253,286
; PRIOR FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 626
; TYPE: PRT
; ORGANISM: *Arachis hypogaea*
US-10-253-286-10

Alignment Scores:
Pred. No.: 2,56e-261 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.6% Indels: 0
DB: 4 Gaps: 0
US-10-728-323-1 (1-2032) x US-10-253-286-10 (1-626)

QY 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTTGCTAGGGATCCTTGTCTCGGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGATGCGCAAGTCATCACCTTACCAGAGAAACACAGAAACCCCTCGGCCAG 169
Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGGTGCTCCAGAGTTGCTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGGTGCACCAAGCTCGAGTATGATCCTCGTGTGTCTATGATCCTCGAGGACACACTGGC 289
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAACGTTCCCTCCAGGGGAGCGGACACGTGGCCGCCCAACCCGGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProProGlyValArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCGTCAACCCCGAGAGAGGAAGGCGCGATGGGACACAGCTGGACCG 409
Db 101 AspAspAspArgArgGlnProArgArgGluGluGlyArgTrpGlyProAlaGlyPro 120
QY 410 AGGAGCGGTGAAGAGAGAGAGACTGGAGACACCAAGAGAGATTGGAGCGGACCAAGT 469

Db 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATCAGCAGCCACGAAAAATAAGCCCGAAGAGAGAGAGAGAAACAAGTAGTGGGAACA 529
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
QY 530 CCAGGTAGCCATGTCAGGGAGAAACATCTCGGAACAACCCCTTCTACTCTCCGCTCAAGG 589
Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCCGCTACGGGAACCAAAACCGGTAGGATCCGGGTCTCTGCAGAGGTTTGAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGTCAAGGCAGTTTCAGAAATCTCCAGAAATCACCGGTATTGTGCAGATCCAGGCCAAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTGTTCTTCCCAGCAGCGCTGATGCTGATAACATCCTCTCTTATCCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGCTGACCGTAGCAAAATGGCAATAACAGAAAGAGCTTTTAACTTTCGAC 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCCATGCACTCAGAAATCCATCCCGTTTCATTCTCTACATCTTGAACCGCATGACAAC 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGCTAAAATCTCCATCCCGCTTAACACACCCCGCCAGTTTGAAGAT 949
Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGAGCAGCCGAGACCAATCATCTCTTCAGCGGCTTCAGCAGGAATACG 1009
Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGAGGCGCCCTTCAATCGGAATTCATGATGATACGAGGGTGTCTGTAGAGAGAAT 1069
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340
QY 1070 GCAGGAGGTGACCAAGAGGAGAGAGCGGAGCGATGAGTACTCGGAGTAGTAGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGGAGTGATAGTCAAAGTGTCAAAGGACACGTTTGAAGAACTTACTAAGCACGCT 1189
Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCGTCTCAAAGAAAGGCTCCGAAAGAGGGAGATATCAACCAACCAATCAACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGNAGAGGCGAGCCCATCTTTTAACTTTGGAGTATTTTTCAGGTGAGAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAAACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAAGGCCATGGTTATCGTCGTCGTCAACAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAACCTGGAACCTTGAACCTCGTGGCTGTAAGAAAGAGCAACAAACAGAGGGGACCGCG 1489
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGlnGlnGlnArgGlyArgArg 480
QY 1490 GAAGAAGAGGAGGACCAAGACCAAGAGAGCGGGAAGTAAACAGAGAGGTGCGTAGGTAC 1549

Db	401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLeuPheGluValLysProAsp	420
QY	1310	AGAAGAAGCCCGAGCTTCAGGACCTGGACATGATGCTACCTGTGTAGAGATCAAGAA	1369
Db	421	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu	440
QY	1370	GGAGCTTTCATGCTCCACACTTCAACTCAAGGCCATGGTTATCTCGTCGTCACAAA	1429
Db	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	460
QY	1430	GGAACTGGAAACCTTGAACCTCGTGGCTGTAAAGAAAAGAGCAACACAGAGGGGACGCGG	1489
Db	461	GlyThrGlyAsnLeuGluLeuAlaValArgLysGluGlnGlnArgGlyArgArg	480
QY	1490	GAAGAAGAGAGACGAAAGACGAAGAGAGAGGGAAGTAACAGAGAGTGCCTAGGTAC	1549
Db	481	GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr	500
QY	1550	ACAGCGAGGTGAAGGAAGCGATGCTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC	1609
Db	501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520
QY	1610	AACGCTTCTCCGAACCTCCACTCTGCTTGGCTTCGGTATCAACGCTCAAAAACACACAGA	1669
Db	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg	540
QY	1670	ATCTTCTTCAGGTGATAAGACAATGTGATAGACAGATAGACGAAGCGAAGGAT	1729
Db	541	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp	560
QY	1730	TTAGCATTCCTCGGTCGGGTGAACAAGTTGAGAGCTCATCAAAAACAGAGAGGAATCT	1789
Db	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	580
QY	1790	CACCTTGTGAGTGCTCGTCTCAATCTCAATCTCAATCTCCGTCGTCCTGAGAAAGAG	1849
Db	581	HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu	600
QY	1850	TCTCCTGAGAAAGAGATCAAGAGGAGGAAACCAAGAGGGGAAGGTCCTCTTTCA	1909
Db	601	SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer	620
QY	1910	ATTTTGAAGCTTTTAAAC	1927
Db	621	IleLeuLysAlaPheAsn	626
RESULT 7			
US-10-899-551-2			
; Sequence 2, Application US/10899551			
; Publication No. US20050063994A1			
; GENERAL INFORMATION:			
; APPLICANT: Caplan, Michael J.			
; APPLICANT: Burks, A. Wesley			
; APPLICANT: Sampson, Hugh A.			
; APPLICANT: Howard, Sosin B.			
; APPLICANT: Bottomly, Kim H.			
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy			
; FILE REFERENCE: 2002834-0233			
; CURRENT APPLICATION NUMBER: US/10/899,551			
; CURRENT FILING DATE: 2004-07-26			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 2			
; LENGTH: 626			
; TYPE: PRT			
; ORGANISM: species Atachis hypogea			
US-10-899-551-2			
Alignment Scores:			
Pred. No.:	2,56e-261	Length:	626
Score:	3286.00	Matches:	626
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0

Query Match:	91.6%	Indels:	0
DB:	5	Gaps:	0
US-10-728-323-1 (1-2032) x US-10-899-551-2 (1-626)			
QY	50	ATGAGAGGAGGAGGTTTCTCCACTGATGCTGTGTCTAGGATCCTTGTCTGGCTTCAGTT	109
Db	1	MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal	20
QY	110	TTCGAACCATGCCAAGTATCATCTTACCAGAAAGAAACAGAAACCCCTGCGCCAG	169
Db	21	SerAlaThrHisAlaLysSerSerProTyrGlnLysThrGluAsnProCysAlaGln	40
QY	170	AGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGCATCCGAGTCT	229
Db	41	ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer	60
QY	230	CGCTGCACCAAGCTCCAGTATGATCTCGTTGTCTATGATCTCCGAGACACACTGGC	289
Db	61	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly	80
QY	290	ACCACCAACCAACGTTTCCCTCCAGGGAGCGGACACGTGGCGCCCAACCCGGAGACTAC	349
Db	81	ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	100
QY	350	GATGATCAGCGCGCTCAACCCCGAGAGAGAGGAGCGCGATGGGAGCCAGCTGCACCG	409
Db	101	AspAspAspArgArgGlnProArgArgGluGluGlyGlyArgTyrProAlaGlyPro	120
QY	410	AGGAGCGTGAAGAGAAAGAAAGACTCGAGACAAACCAAGAGAGATTGGAGCGCACCAAGT	469
Db	121	ArgGluArgGluArgGluGluAspTyrArgGlnProArgGluAspTyrArgArgProSer	140
QY	470	CATCAGCACCCAGGAAATAAGCCGAGAGAGAGAGAGAGAAACAAGAGTGGGAACA	529
Db	141	HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGluTyrGlyThr	160
QY	530	CCAGGTAGCCATGTGAGGAGAAACATCTCGGAACAACCTTCTTACTTCCCGTCAAGG	589
Db	161	ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg	180
QY	590	CGGTTTTAGCACCCGCTACCGGAAACCAAAACCGTAGGATCCGGTCTCGCAGAGGTTTGC	649
Db	181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	200
QY	650	CAAAGTCAAGGCAGTTTCAGAAATCTCCAGAAATCACCGTATTCTGCAGATCGAGGCCAAA	709
Db	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	220
QY	710	CCTAACACTTGTCTTCTCCAGACGCTGATGCTGATAACATCTTGTATTATCCAGCAA	769
Db	221	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	240
QY	770	GGCAAGCCACCGTGACCGTAGCAATGGCAATTAACAGAAAGAGCTTTAATCTTGACGAG	829
Db	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260
QY	830	GGCCATGCATCAGAAATCCATCCCGTTTCATTTCCTACATCTTGAACCCCATGACAAC	889
Db	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280
QY	890	CAGAACTCAGAGTAGCTAAATCTCCATCCCGTTTAAACACACCCGCGCAGTTTGAAGAT	949
Db	281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	300
QY	950	TTCTTCCCGGAGAGCGGAGACCAATCATCTTCTTGTGAGGCGCTTCAGCAGGAATACG	1009
Db	301	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	320
QY	1010	TTGAGCGCCCTTCAATGCGGAATTCATGAGATACGGAGGTTGCTGTTAGAGAGAAT	1069
Db	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn	340

```
QY 1070 GCAGAGGTGAGCAAGAGGAGAGCGGAGCGGATGAGTACTCGAGTACTGAGAAC 1129
Db |||||
QY 341 AlaGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
Db |||||
QY 1130 AATGAGGAGTGTACTCAAGTGTCAAGGACAGGTCAGAACTTACTAAGCAGCT 1189
Db |||||
QY 361 AsnGluGlyValIleValIleValIleValIleValIleValIleValIleValIle 380
Db |||||
QY 1190 AATCCGTCTTCAAGAAAGGCTCCGAGAGAGGAGAGATATACCAACCCCAACTTGTG 1249
Db |||||
QY 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
Db |||||
QY 1250 AGAAGAGGAGCCCGATCTTCTTCAACAATTTGGGAAGTATTGAGGTGAAGCAGAC 1309
Db |||||
QY 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValIleValIle 420
Db |||||
QY 1310 AGAAGAAACCCCGAGCTTCAGGACCTGGACATGCTCACCTGTGTAGAGATCAAGAA 1369
Db |||||
QY 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
Db |||||
QY 1370 GGAGCTTTTGTATGCTCCACACTTCAACTCAAGGCCATGGTTATCTCGTCTCAACAAA 1429
Db |||||
QY 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValIleValIleVal 460
Db |||||
QY 1430 GGAATCGAAACCTTGAATCTCGTGTGTAAGAAAAGACAACAGAGGGGAGCGG 1489
Db |||||
QY 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
Db |||||
QY 1490 GAAGAGAGGAGCAACAGACGAGAGAGGAGGAGGAGTAACAGAGAGGTGCGTAGGTAC 1549
Db |||||
QY 481 GluGluGluGluAspGluAspGluGluGluGluGluGlySerAsnArgGluValArgArgTyr 500
Db |||||
QY 1550 ACAGCGAGGTGAAGGAAGCGCATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCATC 1609
Db |||||
QY 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaHisProValAlaIle 520
Db |||||
QY 1610 AGCGTTCCTCCGAATCCATCTGCTTGGCTTGGTATCAACGCTCAAAACACACAGA 1669
Db |||||
QY 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg 540
Db |||||
QY 1670 ATCTCTCTTCGAGTGATAGGACATCTGATAGACCATAGACAGAGCAAGCGAGGAT 1729
Db |||||
QY 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
Db |||||
QY 1730 TTAGCATTCCTCGGTGCGGTGAACAAAGTTGAGAAGCTCATCAAAACCCAGAGGAATCT 1789
Db |||||
QY 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
Db |||||
QY 1790 CACTTTGTGAGTGCTCGTCTCAATCTCAATCTCAATCTCGTCTCGTCTGAGAAAGAG 1849
Db |||||
QY 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
Db |||||
QY 1850 TCTCCTGAGAAAGAGATCAAGAGGAGGAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
Db |||||
QY 601 SerProGluGlyGluAspGlnGluGluGluGluGluGluGluGluGluGluGluGluSer 620
Db |||||
QY 1910 ATTTTGAAGCTTTTAAAC 1927
Db |||||
QY 621 IleLeuLysAlaPheAsn 626
Db |||||
RESULT 8
US-10-899-551-53
; Sequence 53, Application US/10899551
; Publication No. US20050063994A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Burks, A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Howard, Sosin B.
; APPLICANT: Bottomly, Kim H.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
; FILE REFERENCE: 2002834-0233
; CURRENT APPLICATION NUMBER: US/10/899,551
```

```
; CURRENT FILING DATE: 2004-07-26
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 53
; LENGTH: 635
; TYPE: PRT
; ORGANISM: species Arachis hypogea
US-10-899-551-53

Alignment Scores:
Pred. No.: 1,41e-253 Length: 635
Score: 3192.00 Matches: 605
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.8% Mismatches: 0
Query Match: 89.0% Indels: 0
DB: Gaps: 0

US-10-728-323-1 (1-2032) x US-10-899-551-53 (1-635)
QY 110 TCTGCAACGCATGCCAAGTCATCACTTTACCAAGAAACAGAGAAACCCCTGCGCCAG 169
Db |||||
QY 170 AGTGCCTCCAGAGTTGTCAACAGGAAACCGGATGACTTGAAGCAAAAGCGATGCGAGTCT 229
Db |||||
QY 230 CCGTCCACCAACGCTCGAGTATGATCCTCGTGTGTGTATGATCCTCGAGGACACACTGGC 289
Db |||||
QY 290 ACCACCAACCAACGTTTCCCTCCAGGGAGCGGACACGTCGGCCGCAACCCCGGAGACTAC 349
Db |||||
QY 350 GATGATGACCGCGTCAACCCCGAAGAGAGAGGAGGAGGCGGATGGGAGCCAGCTGACCG 409
Db |||||
QY 410 AGGAGCGTGAAGAGAGAGACACTGGAGACACCAAGAGAGAGATTGGAGGCGACCAAGT 469
Db |||||
QY 470 CATGACGACGACGCAAAATAGGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Db |||||
QY 530 CAGGTAGCCATGTGAGGAGAGAAACATCTCGGAAACACCCCTTCTACTTCCCGTCAAGG 589
Db |||||
QY 590 CGGTTTACCGCTACGGGAACCAAAACCGTAGGATCCGGTCTCTGAGAGGTTTGAC 649
Db |||||
QY 650 CAAAGGTCAAGGCGAGTTTCAGAAATCTCCAGAAATCACCGTATTGTCAGATCCAGGCGCAA 709
Db |||||
QY 710 CTTAACACTTGTGTTCTTCCCAAGCACGCTGATGCTGATATAACATCTTGTGTTTTCAGCAA 769
Db |||||
QY 770 GGGCAAGCCCGCTGACCGTAGCAATTCGCAATTAACAGAAAGAGCTTTTAACTTTCAGCAG 829
Db |||||
QY 830 GGCCATGCACTCAGAAATCCCATCCCGTTTCAATTTCTTCTACATCTTGAACCGCATGACAC 889
Db |||||
QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCCCGTTTAAACACACCCCGGCGAGTTTGAGTAT 949
Db |||||
```

```
Db      277 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 296
QY      950 TTCTTCCCGCGAGCCGAGCAATCATCTTGCAGGCTTCAGCGGAATACG 1009
Db      297 PhePheProAlaSerSerArgAspGlnSerSerTyLeuGlnGlyPheSerArgAsnThr 316
QY      1010 TTGAGGCGCCCTTCAATCGGAAATTCATGACATACGAGGGTGTCTGTAGAGAAAT 1069
Db      317 LeuGluAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 336
QY      1070 GCAGGAGGTGAGCAAGAGAGAGAGCGGAGGCGATGAGTAGTCTCGGAGTAGTGAGAAC 1129
Db      337 AlaGlyGlyGlnGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 356
QY      1130 AATGAAGGAGTGATAGTCAAAAGTGTCAAAGGACACGTTGAAGAACTTACTAAGCACGCT 1189
Db      357 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 376
QY      1190 AATCCGTCTCAAGAAAGGCTCGAAGAGAGGGAGATATCACCAACCCCACTTGTG 1249
Db      377 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 396
QY      1250 AGAAGAGGCGAGCCCATCTTTCAACAACTTTGGGAAGTTATTGAGGTGAAGCCAGAC 1309
Db      397 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 416
QY      1310 AAGAAGAACCCCGCTTCAGGACCTGGACATGATCTCACCTGTGTAGAGATCAAGAA 1369
Db      417 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 436
QY      1370 GAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGGTTATCGTCGTCGTCACAAA 1429
Db      437 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAlaAsnLys 456
QY      1430 GGAACCTGGAACCTTGAACTCGTGCTGTAAAGAAAGAGCAACAACAGAGGGGACCGCG 1489
Db      457 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGlnGlnGlnArgGlyArgArg 476
QY      1490 GAAGAAGAGGAGACGAAGAAGAGAGAGGAGGGAAGTTAACAGAGAGTGCGTAGGTAC 1549
Db      477 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValAlaArgArgTyr 496
QY      1550 ACAGCAGGTTGAAGNAGCGATGTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
Db      497 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 516
QY      1610 AACGCTTCCCTCCCAACTCCCATCTGCTTGGCTTCGGTTATCAACGCTGAAAAACCAACCA 1669
Db      517 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 536
QY      1670 ATCTTCTTCAGGTGATAAGGCAATGTGATAGACCATAGAGAGCAAGCAAGGAT 1729
Db      537 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 556
QY      1730 TTAGCATTCCTCGGTCCGTTGAACAGTTGAGAGCTCATCAAAACCAAGAGGATCT 1789
Db      557 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 576
QY      1790 CACTTTGTGAGTGTCTGCTCCTCAATCTCAATCTCAATCTCCGCTCGCTCCTGAGAAGAG 1849
Db      577 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 596
QY      1850 TCTCTGAGAAAGAGGATCAAGAGGAGGAAACCAAGGAGGAGGATCCACTCCTTTTCA 1909
Db      597 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyLysGlyProLeuLeuSer 616
QY      1910 ATTTTGAAGCTTTTAACT 1927
Db      617 IleLeuLysAlaPheAsn 622
```

RESULT 9

US-09-731-221-78

; Sequence 78, Application US/09731221

```
; Patent No. US20020018778A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael
; TITLE OF INVENTION: Passive Desensitization
; FILE REFERENCE: 2002834-0103
; CURRENT APPLICATION NUMBER: US/09/731,221
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 78
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Arachis
; OTHER INFORMATION: Hypogaea
US-09-731-221-78
```

```
Alignment Scores:
Pred. No.:      4,76e-242      Length:      634
Score:          3052.00      Matches:     584
Percent Similarity: 96.9%      Conservative: 3
Best Local Similarity: 96.4%      Mismatches:  19
Query Match:      85.1%      Indels:      0
DB:               3          Gaps:      0
```

US-10-728-323-1 (1-2032) x US-09-731-221-78 (1-634)

```
QY      110 TCTGCAACCATGCCAAGTCATCACCTTACCAGAGAAACAGAGAACCCCTGCGCCAG 169
Db      16 SerSerThrHisAlaLysSerSerProTyrGlnAlaLysThrGluAsnProCysAlaGln 35
QY      170 AGTGCCTCAGAGTTGTCAAAGAACCGGATGACTTGAAGCAAAAGGATCGGAGTCT 229
Db      36 ArgCysLeuGlnSerCysGlnGlnGluProAspAlaLeuLysGlnLysAlaCysGluSer 55
QY      230 CGCTGCACCAAGCTCGAGTATGATCCTCGTTGTCTATGATCCTCGAGACACACTGGC 289
Db      56 ArgCysThrLysLeuGluTyrAspProArgCysAlaTyrAspProArgGlyHisThrGly 75
QY      290 ACCACCAACCAAGTTTCCCTCTCCAGGGAGCGGACACGTCGGCGCCCAACCCGAGACTAC 349
Db      76 ThrThrAsnGlnArgSerProGlyGluAlaThrArgGlyArgGlnProGlyAspTyr 95
QY      350 GATGATACCGCGCTCAACCCCGAAGAGAGAGGAGCGCGATGGGAGCAGCTGACCG 409
Db      96 AspAspAlaArgArgGlnProArgAlaGluGluGlyGlyArgTrpGlyProAlaGlyPro 115
QY      410 AGGGAGCGTGAAGAGCAAGAGACTGAGACACCAAGAGAGAGATTGGAGGCGACCAAGT 469
Db      116 ArgGluArgGluArgGluGluAspAlaArgGlnProArgGluAspTrpAlaArgProSer 135
QY      470 CATCAGCAGCCCGGAAAAATAAGGCCCGAAGGAGAGAGAGCAAGAGTAGTCGGGAACA 529
Db      136 HisGlnGlnProArgLysAlaArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 155
QY      530 CAGGTAGCCATGTGAGGAGAGAAACATCTCGGAACAACCCCTTCTACTTCCCGTCAAG 589
Db      156 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 175
QY      590 CGGTTTTAGCACCCGCTACCGGAAACCAAAACGGTAGGATCCGGCTCTCGAGAGGTTTGAC 649
Db      176 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 195
QY      650 CAAAGGTCAAGGAGTTTCAGAAATCTCCAGAAATCACCGTATTCGTGAGATCGAGGCCAAA 709
Db      196 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 215
QY      710 CCTAACACTCTTCTTCTTCCCAAGCAGCTGATGCTGATAACATCTCTGTATTATCCAGCAA 769
Db      216 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 235
QY      770 GGGCAAGCCACCGTAGCCGTAGCAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAG 829
```

Db 236 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 255
QY 830 GCCCATGCACCTCAGATCCATCCGCTTCATTTCTACATCTTGAACCGCCATGACAC 889
Db 256 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 275
QY 890 CAGAACTCAGAGTAGTAAATCTCCATGCCCTTAACACACACCCGCCGAGTTGAGGAT 949
Db 276 GlnAsnLeuArgValAlaIleSerMetProValAsnThrProGlyGlnMetGluAsp 295
QY 950 TTCTTCCCGGAGAGCGAGACCAATCATCTTACTTCAGGGCTTCAGCGAGTAATAG 1009
Db 296 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheAlaArgAsnThr 315
QY 1010 TTGGAGCGCCCTTCAATCGGAATTCATAGATACGAGGGTGTCTTAGAAGAGAT 1069
Db 316 LeuGluAlaAlaPheAsnAlaGluAlaAsnGluIleArgValLeuLeuGluAsn 335
QY 1070 GCAGGAGGTGACGACAGAGAGAGCGCAGAGCGATCGAGTACTCGGAGTAGTGAGAAC 1129
Db 336 AlaGlyGlyGluGlnGluAlaArgGlyGlnArgArgTipSerThrArgSerSerGluAsn 355
QY 1130 AATGAAGAGTAGTAAAGTGTCAAAGGACGCTTGAAGAACTTACTTAAGCAGCT 1189
Db 356 AsnGluGlyValIleValIleValIleValSerLysGluHisValGluLeuThrLysHisAla 375
QY 1190 AATCCGTCTCAAGAAAGGCTCCGAGAGAGGGAGATATCACCAACCCCAATCACTTG 1249
Db 376 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProAlaAsnLeu 395
QY 1250 AGAAGAGCGAGCCCATCTTTCTAAACAATTTGGGAAGTTATTTGAGGTGAAGCCAGAC 1309
Db 396 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuAlaGluValLysProAsp 415
QY 1310 AGAAGAAACCCAGCTTCAGACCTGGAATGATGCTCACCTGTGTAGAGATCAAGAA 1369
Db 416 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 435
QY 1370 GGAGCTTTGATGCTCCACACTCAACTCAAGGCCATGGTTATCGTCTGCTCAACAAA 1429
Db 436 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 455
QY 1430 GGAACCTGGAACCTTCAACTCTGGCTGTAAAGAAAAGAGACAAACAGAGGGGACGGCG 1489
Db 456 GlyThrGlyAsnLeuGluLeuValAlaValaLysGlyGlnGlnGlnArgGlyArgArg 475
QY 1490 GAAGAAAGAGGACCGAAGACGAAAGAGGAGGAAAGTAAACAGAGAGTGCCTAGGTAC 1549
Db 476 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgAlaTyr 495
QY 1550 ACAGCGAGTTGAGGAAGCGATGTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
Db 496 ThrAlaArgLeuLysGlyAspValPheIleMetProAlaAlaHisProValAlaIle 515
QY 1610 AACGCTTCTCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTCAAAACCAACACAGA 1669
Db 516 AsnAlaSerSerGluLeuAlaLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 535
QY 1670 ATCTTCTTCGAGGTGATAAGGACAAATGTATAGACCAGATAGAGAAGCAAGGAT 1729
Db 536 IlePheLeuAlaGlyAspAlaAspAsnValIleAspGlnIleGluLysGlnAlaLysAla 555
QY 1730 TTAGCATTCCTGGGTGGGTGNAACAAGTTGAGAACTCATCAAAACCAAGAGGATCT 1789
Db 556 LeuAlaAlaProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 575
QY 1790 CACTTTGTAGTGCTCGTCTCTCAATCTCAATCTCAATCTCCGTCGCTCTCAGAGAAGAG 1849
Db 576 HisPheValAlaAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 595
QY 1850 TCTCTCGAGAAAGAGATCAAGAGGAGGAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAG 1909

Db 596 SerProGluLysGluAspGlnGluGlnGluAsnGlnGlyGlyLysGlyProLeuLeuSer 615
QY 1910 ATTTTGAAGGCTTTTAAAC 1927
Db 616 IleLeuLysAlaPheAsn 621
RESULT 10
US-09-331-631A-21
; Sequence 21, Application US/09331631A
; Patent No. US20020168392A1
; GENERAL INFORMATION:
; APPLICANT: Manners, John M.
; APPLICANT: Marcus, John Paul
; APPLICANT: Goulter, Kenneth C.
; APPLICANT: Green, Jodie L.
; TITLE OF INVENTION: ANTIMICROBIAL PROTEINS
; FILE REFERENCE: CULAN23 001APC
; CURRENT APPLICATION NUMBER: US/09/331,631A
; PRIOR FILING DATE: 1999-06-21
; PRIOR APPLICATION NUMBER: PCT/AU97/00874
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: AU PO 4275
; PRIOR FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Peanut
; US-09-331-631A-21
Alignment Scores:
Pred. No.: 3,78e-241 Length: 614
Score: 3041.00 Matches: 595
Percent Similarity: 96.0% Conservative: 9
Best Local Similarity: 94.6% Mismatches: 7
Query Match: 84.8% Indels: 18
DB: 3 Gaps: 7
US-10-728-323-1 (1-2032) x US-09-331-631A-21 (1-614)

QY 50 ATGAGAGGAGGGTTCCTCCACTGATGCTGTGGTCTAGGATCCTTGTCTCGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGCATGCCAAGTCATCCTTACCAGAGAAACACAGAACCCCTGCGCCAG 169
Db 21 SerAlaThrGlnAlaLys---SerProTyr---ArgLysThrGluAsnProCysAlaGln 38
QY 170 AGGTCCCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCT 229
Db 39 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 58
QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCTTCGAGGACACACTGC 289
Db 59 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly 74
QY 290 ACCACCAACCAAGTTCCCTCCAGGGAGCGACACGTCGGCCGCCAACCCCGAGACTAC 349
Db 75 AlaThrAsnGlnArgHisProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 94
QY 350 GATGATGACCGCGCTCAACCCCGAAGAGAGAGGAGGAGGCGCGATGCGGACAGCTGACCG 409
Db 95 AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGluPro 114
QY 410 AGGAGCGTGAAGAGAAAGAACATGAGAACCAACCAAGAGAAAGATTGGAGGCGACCAAGT 469
Db 115 ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgArgProSer 134
QY 470 CATCAGCAGCCACGGAATAAGGCCCGAAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAG 529
Db 135 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 154

```
QY 530 CCAGGTAGCATGTGAGGGAAGAAACATCTCGGAACAACCCCTTCTACTCCGTCAGG 589
Db 155 ProGlySerGluValArgGluGluThrSerArgAsnAsnProPheTy rPheProSerArg 174
QY 590 CGGTTTACGACCCCGCTACCGGAACCAAAACGGTAGGATCCGGTCTCTGCAGAGTTTGAC 649
Db 175 ArgPheSerThrArgTy rGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 194
QY 650 CAAAGTCAAGGCAGTTTCAGAACTCCAGAACTCACCGTATTGTGTGAGATCGAGGCCAAA 709
Db 195 GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg 214
QY 710 CCTAACACTCTGTTCTTCCAGCACGCTGATGCTGATAACATCTTGTATTCCAGCAA 769
Db 215 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleuValIleGlnGln 234
QY 770 GGGCAAGCCACCGTAGCCGTCAGAAATGGCAATAACAGAAAGAGCTTTAATCTCTGACGAG 829
Db 235 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 254
QY 830 GGCATGCACTCAGAACTCCATCCCGTTTCATTTCTCATCTTGAACCGCCATGACAAC 889
Db 255 GlyHisAlaLeuArgIleProSerGlyPheIleSerTy rIleLeuAsnArgHisAspAsn 274
QY 890 CAGAACTCAGAGTAGCTAAATCTCCATGCCCGTTTAAACACACCCCGCCAGTTTGAGGAT 949
Db 275 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 294
QY 950 TTCTTCCCGCGAGCAGCCAGACCAATCATCTACTTGCAGGCTTCAGCAGGAATACG 1009
Db 295 PhePheProAlaSerArgAspGlnSerSerTy rLeuGlnGlyPheSerArgAsnThr 314
QY 1010 TTGAGGCCCGCTTCAATCGGAATTCATAGATACGAGGGTGTGTTAGAAAGAAT 1069
Db 315 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 334
QY 1070 GCAGGAGGTGACAAAGAGAGAGCGGAGCGGATCGAGTACTCGGAGTAGTGAGAAC 1129
Db 335 AlaGlyGlyGlnGluGluArgGlyGlnArgArgSerThrArgSerSerAsp--- 353
QY 1130 AATGAAGGAGTGATAGTCAAAGTGTCAAAGGACGCTTGAAGAACTTACTAAGCACGCT 1189
Db 354 AsnGluGlyValIleValLysValSerLysGluHisValGlnGluLeuThrLysHisAla 373
QY 1190 AATCCGTTCTCAAAGAAAGCTCCGAAGAAGAGGGAGATATACCAACCCAACTCACTTG 1249
Db 374 LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu 392
QY 1250 AGAAGAGGCGAGCCCGATCTTTCTAAACACTTTGGGAAGTTATTTCAGGTGAGGCCAGAC 1309
Db 393 ArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLysProAsp 412
QY 1310 AAGAAGAACCCCGAGCTTCAGGACCTCGACATGCTCACCTGTGTAGAGATCAAAGAA 1369
Db 413 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 432
QY 1370 GGAGCTTTGATGTCCTCCACACTTCAACTCAAAGCCATGTTATCGTGTGTCACAA 1429
Db 433 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 452
QY 1430 GGAACCTGGAACCTTGAACCTCGTGGCTGTAAGAAAGAGCAACAACAGAGGGCGCCGG 1489
Db 453 GlyThrGlyAsnLeuGluLeuAlaAlaValArgLysGlnGlnGlnArgGlyArgGly 472
QY 1490 GAA-----GAAGAGGAGGACGAAGACGAAGAAGAGGAGGGAAGTAACAGAGGGTG 1540
Db 473 GluGlnGlnLrpGluGluGluGluAspGluGluGluGlySerAsnArgGluVal 492
QY 1541 CGTAGGTACACAGCGAGGTTGAAGGAAGGCGATGTGTTTCATCATGCGCAGAGCTCATCCA 1600
Db 493 ArgArgTy rThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisPro 512
QY 1601 GTAGCCATCAACGCTTCTCCGAACTCCATCTGCTGGCTTCGGTATCAACGCTGAAAC 1660
```

```
Db 513 ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn 532
QY 1661 AACCAAGAGATCTTCTTGACAGTGAAGAGCAATGTGATAGACACAGATAGAGAAGCAA 1720
Db 533 AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln 552
QY 1721 GCGAAGGATTTAGCATTCCTCGGTCCGGTGAACAAGTTGAGAACTCATCAAAACCCAG 1780
Db 553 AlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGln 572
QY 1781 AAGGAATCTCACTTTGTGAGTGCTCGTCTCTCAATCTCAATCTCAATCTCCGCTCTCCT 1840
Db 573 ArgGluSerHisPheValSerAlaArgProGlnSerGlnSerProSer----- 588
QY 1841 GAGAAAGAGTCTCTCGAGAAAGAGATCAAGAGGAGGAGAAACCAAGAGGAGGAGGTCCA 1900
Db 589 -----SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyPro 605
QY 1901 CTCCTTTCATTTTGAAGGCTTTTAAC 1927
Db 606 LeuLeuSerIleLeuLysAlaPheAsn 614

RESULT 11
US-09-847-208-27
; Sequence 27, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
; TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES
; FILE REFERENCE: UC67.002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Arachis hypogaea (Peanut)
US-09-847-208-27

Alignment Scores:
Pred. No.: 3,78e-241 Length: 614
Score: 3041.00 Matches: 595
Percent Similarity: 96.0% Conservative: 9
Best Local Similarity: 94.6% Mismatches: 7
Query Match: 84.8% Indels: 18
DB: 3 Gaps: 7

US-10-728-323-1 (1-2032) x US-09-847-208-27 (1-614)
QY 50 ATGAGAGGAGGGGTTTCTCCACTGATGCTGTGTAGGATCTCTGTCTCGTGGTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGCATGCCAAGTCATCACCTTACCAGAGAAACACAGAGAACCCCTCGCCCCAG 169
Db 21 SerAlaThrGlnAlaLys---SerProTy r---ArgLysThrGluAsnProCysAlaGln 38
QY 170 AGGTGCTTCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAAGGCATCGAGTCT 229
Db 39 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 58
QY 230 CGCTGCACCAAGCTCCAGTATGATCTCGTTGTGTCTATCATCTCTCGAGGACACACTGGC 289
Db 59 ArgCysThrLysLeuGluTy rAspProArgCysValTy rAsp-----ThrGly 74
QY 290 ACCACCAACCAAGTTTCCCTCCAGGGAGCGGACACGTGGCGGCCCAACCCGAGAGACTAC 349
Db 75 AlaThrAsnGlnArgHisProProGlyGluArgThrArgGlyArgGlnProGlyAspTy r 94
```

QY 350 GATGATGACCGCGCTCAACCCGAGAGAGAGAGGCGCGATGGGACACGAGCTGGACCG 409
DB |||||
95 AspAspAspArgArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGluPro 114
QY 410 AGGAGCGGTAAAGAGAAGAAGACTGGAGACCAACCAAGAGAGATTGGAGGCGCAACAGT 469
DB |||||
115 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 134
QY 470 CATCACACCGCCGGAATATAGGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
DB |||||
135 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGlnTrpGlyThr 154
QY 530 CAGGTAGCCATGTGAGGGAAGAAACATCTCGGAAACAACCTTTTCTACTTCCCGTCAAGG 589
DB |||||
155 ProGlySerGluValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 174
QY 590 CGGTTTAGCACCGCTACGGGAACCAAAACGGTAGGATCCGGGTCTCGAGAGGTTTGAC 649
DB |||||
175 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 194
QY 650 CAAAGGTCAAGGCAGTTTCAGAACTCCAGAACTCCAGATTCACCGTATTCTGCAGATCCGAGGCCAAA 709
DB |||||
195 GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg 214
QY 710 CCTAAACACTCTTCTTCCCAAGCACGCTGTATGCTGATAAACAATCTCTTGTATCCAGCAA 769
DB |||||
215 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 234
QY 770 GGGCAAGCCACCGTGACCGTAGCAATGCAATACAGAAAGAGCTTTAATCTTGACGAG 829
DB |||||
235 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 254
QY 830 GCCCATGCACTCAGATCCATCCGCTTTCATTTCTTACATCTTGACCGCCATGACAC 889
DB |||||
255 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 274
QY 890 CAGAACTCAGAGTACGTAATCTCCATGCCCGTAAACACACCCCGCCGCTTTCAGGAT 949
DB |||||
275 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 294
QY 950 TTCTTCCCGGAGCAGCGAGACCAATCATCTCTTTCAGCGCTTCAGCAGGAAATACG 1009
DB |||||
295 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 314
QY 1010 TTGGAGCGCCCTTCAATCGGAATTCATGATACGAGCGGCTGCTTAGACAGAT 1069
DB |||||
315 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 334
QY 1070 GCAGGAGGTGAGCAAGAGAGAGCGGAGCGATGGAGTACTCGGAGTAGTGAGAAC 1129
DB |||||
335 AlaGlyGlyGluGlnGluArgGlyGlnArgArgSerThrArgSerSerAsp--- 353
QY 1130 AATGAAGAGTAGTATCTAAAGTGTCAAGAGACAGCTTGAAGAACTTACTAAGCAGCT 1189
DB |||||
354 AsnGluGlyValIleValLysValSerLysGluHisValGlnGluLeuThrLysHisAla 373
QY 1190 AATCCGTCTCAAGAAAGCTCCGAGAGAGGAGATACCAACCCCACTCACTTG 1249
DB |||||
374 LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu 392
QY 1250 AGAAGAGCGGAGCGCCGATCTTTCTAAACAATTTGGGAAGTTATTTCAGGTGAAGCCAGAC 1309
DB |||||
393 ArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLysProAsp 412
QY 1310 AAGAAGAACCCCGCTTCAGGACCTGGACATGATCTCACCTGTGTAGAGATCAAGAA 1369
DB |||||
413 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 432
QY 1370 GAGGCTTTGATGCTCCCACTTCAACTCAAGGCCATGGTTATCGTCTCGTCAACAA 1429
DB |||||
433 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 452

QY 1430 GGAACCTGAAACCTTTGAATCTCGTGGCTGTAAAGAAAGAGCAACAACAGAGGGGACGGCGG 1489
DB |||||
453 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 472
QY 1490 GAA-----GAAGAGGAGACCAAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1540
DB |||||
473 GluGlnGluTrpGluGluGluAspGluGluGluGluGluGluGluGluGluVal 492
QY 1541 CQTAGGTACACAGCCAGGTGGAAGGAGCGATGTGTTTCATCATGCCAGCAGCTCATCCA 1600
DB |||||
493 ArgArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisPro 512
QY 1601 STAGCCATCAACGCTTCTCCGAACCTCATCTGCTTGGCTTCGGTATCAACCTGAAAAAC 1660
DB |||||
513 ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn 532
QY 1661 AACCAACAGAACTTCTTCAGGTGATTAAGNCAATGTGATAGACCATAGAGAGCAA 1720
DB |||||
533 AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln 552
QY 1721 GCGAAGGATTTAGCATTCCTCGGTGGGTGAACAAGTTGAGAAGCTCATCAAAACCCAG 1780
DB |||||
553 AlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGln 572
QY 1781 AAGGAATCTCACTTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCGTCTCTCT 1840
DB |||||
573 ArgGluSerHisPheValSerAlaArgProGlnSerGlnSerProSer----- 588
QY 1841 GAGAAGAGTCTCTCGAAGAGATCAAGAGGAGGAGAAACCAAGGAGGAGGAGGAGGAGGAG 1900
DB |||||
589 -----SerProGluLysGluAspGlnGluGluGluGluGluGluGluGlyGlyGlyPro 605
QY 1901 CTCCTTTCAATTTTGAAGGCTTTTAAAC 1927
DB |||||
606 LeuLeuSerIleLeuLysAlaPheAsn 614

RESULT 12

US-10-147-095-21
; Sequence 21, Application US/10147095
; Publication No. US20030171274A1
; GENERAL INFORMATION:
; APPLICANT: Manners, John M.
; APPLICANT: Marcus, John Paul
; APPLICANT: Goulter, Kenneth C.
; APPLICANT: Green, Jodie L.
; TITLE OF INVENTION: ANTIMICROBIAL PROTEINS
; FILE REFERENCE: CULN23.001APC
; CURRENT APPLICATION NUMBER: US/10/147,095
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US/09/331,631A
; PRIOR FILING DATE: 1999-06-21
; PRIOR APPLICATION NUMBER: PCT/AU97/00874
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: AU PO 4275
; PRIOR FILING DATE: 1996-12-20
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Peanut
US-10-147-095-21

Alignment Scores:
Pred. No.: 3,78e-241 Length: 614
Score: 3041.00 Matches: 595
Percent Similarity: 96.0% Conservative: 9
Best Local Similarity: 94.6% Mismatches: 7
Query Match: 84.8% Indels: 18
DB: 4 Gaps: 7

US-10-728-323-1 (1-2032) x US-10-147-095-21 (1-614)

QY 50 ATGAGAGGAGGCTTTCTCCACTGATCCTGTTCTAGGATCCTGCTGCTTCAAGTT 109
Db |||||
1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGATCCCAAGTCATCACCTTACCAAGAAAGAAACAGAGAACCCCTGCGCCAC 169
Db ||||| :|||
21 SerAlaThrGlnAlaLys---SerProTyr--ArgLysThrGluAsnProCysAlaGln 38
QY 170 AGTGTCTCAGAGTTGTCAACAGGAAACCGATGACTTGAAGCAAAAGGATCGAGTCT 229
Db |||||
39 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 58
QY 230 CGTGCACCAAGCTCAGATATGATCCTCTGTGTCTATGATCCTCGAGGACACATGCG 289
Db |||||
59 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly 74
QY 290 ACCACCAACCAACGTTCCCTCCAGGGAGCGGACACAGTGGCGGCCCAACCCGAGACTAC 349
Db |||||
75 AlaThrAsnGlnArgHisProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 94
QY 350 GATGATGACCGCTCAACCCCAAGAGAGAAAGGAGCGCGATGGGACCAAGCTGACCG 409
Db |||||
95 AspAspArgArgGlnProArgGluGluAspTrpArgGluAspTrpArgProSer 134
QY 470 CATCAGCAGCCAGGAAATTAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Db |||||
135 HisGlnGlnProArgLysIleArgProGluGluArgGluGluGluGluGluGluGluGlu 154
QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAAACCTTTTACTTCCCGTCAAGG 589
Db |||||
155 ProGlySerGluValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 174
QY 590 CGGTTAGCACCGCTACGGGACCAACAAACGGGTAGGATCCGGCTCTGCAGAGGTTGAC 649
Db |||||
175 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 194
QY 650 CAAAGGTCAAGGAGGATTTCAAGATCTCCAGAATCACCGTATTGTGCAGATCGAGGCCAA 709
Db ||||| :|||
195 GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaArg 214
QY 710 CCTAACACTTGTCTTCTCCAGCAGCAGCTGATGCTGATAACATCTCTGTTATTCACGCAA 769
Db |||||
215 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 234
QY 770 GGGCAACCCACCGTAGCCGTAGCAATGGCAATAACAGAAAGAGCTTTATCTTGACGAG 829
Db |||||
235 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 254
QY 830 GGCCATGCATCAGAATCCCATCCGTTTCATTTCTTACATCTTTGAACCCGCCATGACAAC 889
Db |||||
255 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 274
QY 890 CAGAACCTCAGATAGTAAATCTCCATGCCCGCTTAAACACACCCGCGCCAGTTTGAGGAT 949
Db |||||
275 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 294
QY 950 TTCTTCCCGGAGCAGCCGAGCAATCATCTTACTTTCAGGCGCTTCAGCAGCAATACG 1009
Db |||||
295 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 314
QY 1010 TTGAGGCGCGCTTCAATGCGGAATTCATAGATACGAGGGGTGCTGTTTAGAAGAGAAT 1069
Db |||||
315 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 334
QY 1070 GCAGGAGGTAGCAAGAGGAGAGAGGCGCAGGCGATGGAGTACTCGGAGTAGTGAGAAC 1129
Db ||||| :|||
335 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgSerThrArgSerSerAsp--- 353
QY 1130 AATGAAGGAGTGATAGTCAAAAGTGTCAAAGGAGCAGCCTTGAGAACTTACTAAGCAGCT 1189

Db ||||| :|||
354 AsnGluGlyValIleValLysValSerLysGluHisValGlnGluLeuThrLysHisAla 373
QY 1190 AAATCCGTCTCAAGAAAGGCTCCGAAGAAGAGGAGATATCACCAACCCAACTCAACTTG 1249
Db |||||
374 LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu 392
QY 1250 AGAGAAGGCGAGCCCGATCTTTCTAACACTTTGGGAAGTTATTTGAGGTGAGCGAGAC 1309
Db ||||| :|||
393 ArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLysProAsp 412
QY 1310 AAGAAGAACCCCGAGCTTCAGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
Db |||||
413 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 432
QY 1370 GGAGCTTTGATGCTCCACACTTCACATCAAAAGGCATGGTTATTCGTCTGCTCAACAA 1429
Db |||||
433 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 452
QY 1430 GGAACCTGGAACCTTGAACCTCGTGTAAAGAAAGACCAACACAGAGGGGAGCGCG 1489
Db |||||
453 GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnArgGlyArgArg 472
QY 1490 GAA-----GAAGAGGAGGACGAACGACGAAGAGAGGAGGAAGTAAACAGAGGTTG 1540
Db ||||| :|||
473 GluGlnGluTrpGluGluGluGluAspGluGluGluGlySerAsnArgGluVal 492
QY 1541 CTTAGGTACACAGCAGGTTGAAGAGCGCATGTGTTTCATCATGCCAGCAGCTCATCCA 1600
Db |||||
493 ArgArgTyrThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaHisPro 512
QY 1601 CTAGCCATCAACGCTTCTCCGAACCTCATCTGCTTGGCTTCGGTATCAACCTGAAAAAC 1660
Db |||||
513 ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn 532
QY 1661 AACCACAGAAATCTCTTCGAGTGATAGGACAATGTGATAGACCAGATAGAGAGCAA 1720
Db |||||
533 AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln 552
QY 1721 GCGAAGGATTAGATTCCTCGGTGCGGTGAACAAAGTTGAGAAGCTCATCAAAACCCAG 1780
Db |||||
553 AlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGln 572
QY 1781 AAGGAATCTCACTTTGTGAGTCTGCTCTCATCTCAATCTCAATCTCAATCTCGTCTCTCT 1840
Db ||||| :|||
573 ArgGluSerHisPheValSerAlaArgProGlnSerGlnSerProSer----- 588
QY 1841 GAGAAAGAGTCTCTCAGAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1900
Db |||||
589 -----SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyPro 605
QY 1901 CTCCTTTCAATTTGAAGCTTTTAAAC 1927
Db |||||
606 LeuLeuSerIleLeuLysAlaPheAsn 614

RESULT 13

US-10-100-303A-8

; Sequence 8, Application US/10100303A

; Publication No. US20030202980A1

; GENERAL INFORMATION:

; APPLICANT: Caplan, et al.

; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction

; TITLE OF INVENTION: to Allergy

; FILE REFERENCE: 2002834-0166

; CURRENT APPLICATION NUMBER: US/10/100,303A

; CURRENT FILING DATE: 2002-03-18

; NUMBER OF SEQ ID NOS: 138

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 8

; LENGTH: 614

; TYPE: PRF

; ORGANISM: Arachis hypogaea, Prot/(Nucleo-Ara h 1

US-10-100-303A-8

Alignment Scores: 3.78e-241 Length: 614
 Pred. No.: 3041.00 Matches: 595
 Score: 96.0% Conservatives: 9
 Percent Similarity: 94.6% Mismatches: 7
 Best Local Similarity: 94.6% Indels: 18
 Query Match: 84.8% Gaps: 7
 DB: 4

US-10-728-323-1 (1-2032) x US-10-100-303A-8 (1-614)

QY 50 ATGAGAGGAGGTTTCTCCACTGATGCTGTGGTCTAGGATCCTTGTCTCGCTTCAGTT 109
 |||||
 Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
 |||||

QY 110 TCTGCAACCGCATGCCAAGTCATCCTTACCAGAGAACAGAAACAGAGAACCCCTGCCCCAG 169
 |||||
 Db 21 SerAlaThrGlnAlaLys---SerProTyr---ArgLysThrGlnAsnProCysAlaGln 38
 |||||

QY 170 AGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCT 229
 |||||
 Db 39 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 58
 |||||

QY 230 CGGTGACCAAGCTCAGTATGATCTCTGTGTCTATGATCTCTCGAGGACACACTGGC 289
 |||||
 Db 59 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAsp-----ThrGly 74
 |||||

QY 290 ACCACCAACCAAGTTCCTCCAGGGAGCGGACAGCTGGCGCCCAACCCGGAGACTAC 349
 |||||
 Db 75 AlaThrAsnGlnArgHisProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 94
 |||||

QY 350 GATGATGACCGCGCTCAACCCCGAAGAGAGAGGCGCGATGGGACCGAGCTGGACCG 409
 |||||
 Db 95 AspAspAspArgArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGluPro 114
 |||||

QY 410 AGGGAGCGTGAAGAGAAGAAGACTGGAGACAACCAAGAGAAGATTGGAGGGCACAAGT 469
 |||||
 Db 115 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 134
 |||||

QY 470 CATCAGCAGCCAGGAATAGGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
 |||||
 Db 135 HsGlnGlnProArgLysIleArgProGluGlyArgGlyGluGlnGluGlnGluTyrGlyThr 154
 |||||

QY 530 CCAGGTAGCATGTGAGGGAAGAAACATCTCGGAACAAACCTTTCTACTTCCCGTCAAGG 589
 |||||
 Db 155 ProGlySerGluValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 174
 |||||

QY 590 CGGTTTAGCACCCGCTACGGGAACCAAAACGGTAGATCCGGGTCTCGAGAGTTTGAC 649
 |||||
 Db 175 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 194
 |||||

QY 650 CAAGGTCAGGCAGTTTCAGAAATCTCCAGAAATCCCGTATTGTGCAGATCCGAGGCCAAA 709
 |||||
 Db 195 GlnArgSerLysGlnPheGlnAsnLeuGlnAsnHisArgGileValGlnIleGluAlaArg 214
 |||||

QY 710 CCTAACACTCTGTCTTCTCCCAAGCACGCTGATGCTGATAACATCTTGTATCCAGCAA 769
 |||||
 Db 215 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 234
 |||||

QY 770 GGGCAAGCCACCGTAGCCCTAGCAATATGCAATATACAGAAAGAGCTTAACTTTCAGCAG 829
 |||||
 Db 235 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 254
 |||||

QY 830 GGCCTAGCACTCAGAAATCCATCCGTTTCATTCTCATCATCTTGAACCGCCATCACAC 889
 |||||
 Db 255 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 274
 |||||

QY 890 CAGAACTCAGAGTAGTAAAATCTCCATGCCCGTTAACACACCCCGCCAGTTTGGAGAT 949
 |||||
 Db 275 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 294
 |||||

QY 950 TTCCTCCCGGAGCAGCCGAGACCAATCATCTTCTACTTCGAGGCTTCAGCAGGAATACG 1009
 |||||

RESULT 14

US-10-100-303A-55

; Sequence 55, Application US/10100303A

; Publication No. US20030202980A1

; GENERAL INFORMATION:

; APPLICANT: Caplan, et al.

; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction

Db 295 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 314
 QY 1010 TTGAGGCGCGCTTCAATCGCAATTCATGAGATACGAGGCTCTCTGTAGAAGAGAAAT 1069
 |||||
 Db 315 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn 334
 |||||

QY 1070 GCAGAGGTGAGCAAGAGAGAGAGCGGAGGCGATCGAGTACTCGGAGTAGTGAGAAC 1129
 |||||
 Db 335 AlaGlyGlyGluGlnGluArgGlyGlnArgArgSerThrArgSerSerAsp--- 353
 |||||

QY 1130 AATGAGGAGTGTAGTCAAAAGTCTCAAGAGGACGCTTGAAGACTTACTTAAGCACGCT 1189
 |||||
 Db 354 AsnGluGlyValIleValLysValSerLysGluHisValGlnGluLeuThrLysHisAla 373
 |||||

QY 1190 AAATCCGCTCAAGAAAGGCTCCGAAGAGAGGAGATATCACCAACCCCAATCAACTTG 1249
 |||||
 Db 374 LysSerValSerLysLysGlySerGluGluGlu---AspIleThrAsnProIleAsnLeu 392
 |||||

QY 1250 AGAGAAAGCGAGCCGATCTTTCTTAACTTTGGAAAGTTATTGTAGGTGAAGCCAGAC 1309
 |||||
 Db 393 ArgAspGlyGluProAspLeuSerAsnAsnPheGlyArgLeuPheGluValLysProAsp 412
 |||||

QY 1310 AAGAAGAACCCCGAGCTTCAGGACCTGCACATGATCTCACCTGTGTAGAGATCAAGAA 1369
 |||||
 Db 413 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 432
 |||||

QY 1370 GGAGCTTTGATGTCTCCACACTTCAACTCAAGGCGCATGTTATCTCGTCTCGTCAACAAA 1429
 |||||
 Db 433 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 452
 |||||

QY 1430 GGAATCGAAACCTTGAACTCGTGGCTGTAAAGAAAGCAACAAACAGAGGGGACGGCG 1489
 |||||
 Db 453 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArg 472
 |||||

QY 1490 GAA-----GAAGAGGAGCAGAACAGCAAGAGAGGGGAGTAACAGAGAGCTG 1540
 |||||
 Db 473 GluGlnGluTrpGluGluGluGluAspGluGluGluGluGlySerAsnArgGluVal 492
 |||||

QY 1541 CGTAGTACACAGCGAGGTTCAAGGAGCGCATGTGTTTCATCATCCAGCAGCTCATCCA 1600
 |||||
 Db 493 ArgArgTyrThrAlaArgLeuLysGluLysValPheIleMetProAlaAlaHisPro 512
 |||||

QY 1601 GTAGCCATCAACGCTTCTCCGAATCCATCTGTTGGCTTCGGTATCAACGCTCAAAAC 1660
 |||||
 Db 513 ValAlaIleAsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsn 532
 |||||

QY 1661 AACCAACAGAACTTCTTGCAGGTCATAGACCAATGTGTAGACCATAGACATAGAGAGCAA 1720
 |||||
 Db 533 AsnHisArgIlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGln 552
 |||||

QY 1721 GCGAAGGATTTAGCATTCCTCGGTCCGGTGAACAAAGTTGAGAACTCATCAAAACCCAG 1780
 |||||
 Db 553 AlaLysAspLeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGln 572
 |||||

QY 1781 AAGGAATCTCACTTTGTGAGTGTCTCTCAATCTCAATCTCAATCTCCGCTCGTCTCT 1840
 |||||
 Db 573 ArgGluSerHisPheValSerAlaArgProGlnSerGlnSerProSer----- 588
 |||||

QY 1841 GAGAAAGAGTCTCTGAGAAAGAGGATCAAGAGGAGGAAACCAAGAGGAGGAGGAGGTCCA 1900
 |||||
 Db 589 -----SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyPro 605
 |||||

QY 1901 CTCCTTTTCAATTTTGAAGGCTTTTAAAC 1927
 |||||
 Db 606 LeuLeuSerIleLeuLysAlaPheAsn 614
 |||||

THIS PAGE LEFT BLANK

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 21:23:02 ; Search time 60.8696 Seconds

(without alignments)
4400.310 Million cell updates/sec

Title: US-10-728-323-1

Perfect score: 3586

Sequence: 1 aataatcatatattcatc.....cgttgtgctgtttcttccc 2032

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/abs/ABSSWEB spool/US10728323/runat_15052006.172130.22366/app_query.fasta.1
-DB=A_Geneseq -QFWT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -
UNITS=bits -START=1 -ENDS=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p
-USER=US10728323 @CGN.1.1.605 @runat_15052006.172130.22366 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: A_Geneseq 21:*
- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*
- 9: Geneseq2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3286	91.6	626	2 AAY15244	Aay15244 Peanut al
2	3286	91.6	626	2 AAY25657	Aay25657 Peanut al
3	3286	91.6	626	6 ABUS2412	Abus2412 Peanut Ar
4	3286	91.6	626	7 ADC34906	Adc34906 Peanut al
5	3286	91.6	626	9 ADV97601	Adv97601 Peanut Ar
6	3282	91.5	626	6 AAU04706	Aau04706 Anaphylac
7	3282	91.5	626	6 ABUS2570	Abus2570 Peanut Ar
8	3280	91.5	626	6 ABUS2568	Abus2568 Peanut Ar
9	3280	91.5	626	6 ABUS2571	Abus2571 Peanut Ar

10	3280	91.5	626	6 ABUS2574	Abus2574 Peanut Ar
11	3280	91.5	626	6 ABUS2573	Abus2573 Peanut Ar
12	3279	91.4	626	8 ADM12053	Adm12053 Arachis h
13	3278	91.4	626	6 ABUS2569	Abus2569 Peanut Ar
14	3278	91.4	626	6 ABUS2572	Abus2572 Peanut Ar
15	3274	91.3	626	2 AAU22150	Aau22150 Peanut al
16	3269.5	91.2	625	7 ADG27464	Adg27464 Peanut Ar
17	3251	90.7	626	8 ADO38271	Ado38271 Peanut al
18	3052	85.1	634	3 AAB33599	Aab33599 Modified
19	3052	85.1	634	4 AAU04709	Aau04709 Modified
20	3052	85.1	634	4 AAU05034	Aau05034 Modified
21	3041	84.8	614	2 AAU22149	Aau22149 Peanut al
22	3041	84.8	614	2 AAU22149	Aau22149 Peanut al
23	3041	84.8	614	7 AAU52413	Aau52413 Peanut Ar
24	3041	84.8	614	7 ADG27465	Adg27465 Peanut Ar
25	3034	84.6	614	9 ADV97606	Adv97606 Peanut Ar
26	2130	59.4	415	4 AAU40913	Aau40913 Ara h 1 a
27	1700	47.4	335	9 ADV97604	Adv97604 Peanut Ar
28	1281	35.7	605	2 AAU62838	Aau62838 Glycine m
29	1281	35.7	605	7 ADG27564	Adg27564 Soybean B
30	1276	35.6	605	2 AAU40999	Aau40999 Soybean b
31	1274.5	35.5	623	8 ADT58763	Adt58763 Plant pol
32	1190	33.2	559	5 AEG71270	Aeg71270 Glycine m
33	1180.5	32.9	543	5 AEG71271	Aeg71271 Glycine m
34	1165	32.5	417	5 AEG71269	Aeg71269 Glycine m
35	1157	32.3	417	6 ABP97239	Abp97239 Mature be
36	1157	32.3	425	6 ABP97238	Abp97238 FLAG-tag
37	1156	32.2	439	6 ABP97235	Abp97235 Glycine m
38	1117.5	31.2	390	6 ABP97236	Abp97236 Glycine m
39	874	24.4	324	8 ADX75042	Adx75042 Plant ful
40	867	24.2	344	8 ADX89039	Adx89039 Plant ful
41	865.5	24.1	666	2 AAU62829	Aau62829 Macadamia
42	863.5	24.1	625	2 AAU62830	Aau62830 Macadamia
43	849.5	23.7	666	2 AAU62828	Aau62828 Macadamia
44	828	23.1	316	8 ADX79896	Adx79896 Plant ful
45	828	23.1	316	8 ADX74539	Adx74539 Plant ful

ALIGNMENTS

RESULT 1

AAY15244

ID AAY15244 standard; protein; 626 AA.

XX AC AAY15244;

XX AC AAY15244;

DT 17-OCT-2003 (revised)

DT 09-NOV-1999 (first entry)

XX

DE Peanut allergen, Ara h 1, amino acid sequence.

XX allergy; immune response; transgenic; allergen; epitope;

KW immunoglobulin E; Ig E; binding site; peanut.

XX Arachis hypogaea.

OS Arachis hypogaea.

PN WO9938978-A1.

XX

PD 05-AUG-1999.

XX

PF 29-JAN-1999; 99WO-US002031.

XX

PR 31-JAN-1998; 98US-0073283P.

PR 13-FEB-1998; 98US-0074590P.

PR 13-FEB-1998; 98US-0074624P.

PR 13-FEB-1998; 98US-0074633P.

XX 27-AUG-1999; 98US-00141220.

(UYAR-) UNIV ARKANSAS.

PA (UYNY) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.

XX (SOSI/) SOSIN H.

PI Sosin H, Bannon GA, Burks AW, Sampson HA;

XX WPI; 1999-479189/40.
DR N-PSDB; AA206382.
XX Modified allergen with reduced IgE binding, useful for treating e.g. allergies.
PT
PT
XX
PS Disclosure; Page 35-37; 46pp; English.

XX This is the amino acid sequence of the Ara h 1 protein from Arachis hypogaea. The Ara h 1 protein has 23 IgE (immunoglobulin E) binding epitopes, four of which are immunodominant (AAV15247, AAV15249, AAV15250 and AAV15263). By modifying the IgE binding sites the ability of the allergen to provoke an immune response is downregulated. The epitopes of the IgE binding sites can therefore be modified in genetically engineered plants and animals to elicit less of an allergic response. (Updated on 17 -OCT-2003 to standardise OS field)

XX Sequence 626 AA;

Alignment Scores:
Pred. No.: 1.59e-304 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.6% Indels: 0
DB: 2 Gaps: 0

US-10-728-323-1 (1-2032) x AAV15244 (1-626)

QY 50 ATGAGAGGAGGTTTCTCCACTGATGCTGTTCTAGGATCTCTGCTGGCTTCAGTT 109
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCACGATGCCAAGTCAATCACCTTACAGAGAAACAGAGAACCCCTCGCCAG 169
DB 21 SerAlaThrHisAlaLysSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCT 229
DB 41 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGGTGCACCAAGCTCGAGTATGATCTCGTTGTCTATGATCTCGAGGACACACTGGC 289
DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAACGTTCCCTCCAGGGGAGGACACGTCGCGCCCAACCCGGAGACTAC 349
DB 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCGTCAACCCGAGAGAGAGAGGCGCGATGGGACCCAGCTGACCG 409
DB 101 AspAspAspArgGlnProArgArgGluGluGlyArgTyrGlyProAlaGlyPro 120
QY 410 AGGAGCGTCAAGAGAGAGAGCTGGAGACACCAAGAGAGATTGGAGCGACCAAGT 469
DB 121 ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATCAGCAGCCACGAAATAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
QY 530 CCAGGTAGCCATGTGAGGGAGAAACATCTCGGAACCAACCTTTCTACTTCCCGTCAAGG 589
DB 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCCGCTACGGGACCAAAACGGTAGGATCCGGTCCCTGACAGAGTTTGAC 649
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGGTCAAGGCGATTTCAAGATCTCCAGAATCACCGTATTTGTGAGATCGAGGCCAAA 709
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220

QY 710 CCTAACACTCTTGTCTTCCCAAGCACGCTGATGCTGATAACATCCTTCTTATCCAGCA 769
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTACCGTAGCAATGCAATACAGAAAGAGCTTTTAACTTTGACGAG 829
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCCATGCACTCAGAAATCCCTCCGTTTCATTTCTACATCTTGAACCCCATGCAAC 889
DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCCGCTTAACACACCCGCCAGTTGAGGAT 949
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGGAGCAGCAGCAGCAATCATCTACTTCAGCGCTTCAGCAGGAATACG 1009
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGGCGCTTCAATCGGAATTCATAGATACGAGGGTGTCTTTAGAGAGAT 1069
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGAGAGAGCGGACGATGAGTACTCGGAGTAGTGAGAAC 1129
DB 341 AlaGlyGlyGluGlnGluArgGlyGlnArgTyrTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGAGTAGTAGTCAAGTGTCAAAGAGACGTTTGAAGAACTTACTAAGCACGT 1189
DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCGCTCTAAAGAAAGGCTCCGAAAGAGGGAGATATCACCAACCAATCAACTTG 1249
DB 381 LysSerValSerLysLysGlySerGluGluGluArgAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAGGCGAGCCCGATCTTCTTAACAATTTGGGAAGTATTTCAGGTGAGCGAC 1309
DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AGAAGAAACCCCGAGCTTCAGGACCTCGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATCGTTATCTGCTGCTCAACAA 1429
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAACCTGGAACCTTCAACTCGTGGCTGAAGAAAAGAGACCAACAGAGGGGACCGCG 1489
DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAGAGAGGAGACCAAGACGAAGAGAGGGAGGTAAACAGAGAGGTGCTAGGTAC 1549
DB 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCAGGTTGAAGAGGCGATGTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAATCCATCTGCTTGGCTTCGGTATCAACGCTCAAAACCAACACCA 1669
DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTCGAGGTGATAGGACAAATGTGATAGACCAGATAGAGAGCAAGCAAGGAT 1729
DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTGAGCAAGTTTGAGAACTCATCAAAACCAACGAGAGGATCT 1789
DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580

Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluLeuArgValLeuLeuGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGAGAGAGGCGAGGCGATGAGTACTCGAGTAGTGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerGluAsn 360
QY 1130 AATGAAGGAGTGTAGTCAAAAGTGTCAAAGGACGCTTGAAGAACTTACTAAGCAGCT 1189
Db 361 AsnGluGlyValIleValIleValSerLysGluHisValGluLeuThrLysHisAla 380
QY 1190 AATCGCTCTCAAGAAGGCTCCGAAGAGAGGAGATATCACCAACCCATCACTTGG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAAGGCGCCGATCTTTCTAAACACTTTGGGAAGTTATTTCAGGTGAAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAACCCCGAGCTTCAGGACCTGGACATGCTCACCTGTGTAGAGATCAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTGTGCTCCACACTTCAACTCAAGGCCATGGTTATCGTTCGTCACAAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAACCTGGAACCTTGAACCTCGTGGCTGTAAAGAAAGAGCAACAACAGAGGCGCGG 1489
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArg 480
QY 1490 GAAGAAGAGGAGCAGAACGAAAGAGAGGAGGAGTAACAGAGAGTGCTAGGTAC 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTGAAGAGGCGATGTCTCATCATCCAGCAGCTCATCCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AAGCTTCTCCGAACCTCCACTCTGCTTGGCTTCGCTATCAACGCTGAAACACACACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCCTCCAGTGATAGGCAATGTGATAGACCATAGAGAGCAAGCAAGGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTGGGTGAACAGTTGAGAGCTCATCAAAACCAAGAGGAATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValIleGluLysLeuLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTGCTCGCTCCTCAATCTCAATCTCAATCTCCGCTCGTCTCCTGAGAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerProGluLysGlu 600
QY 1850 TCTCTGAGAAAAGAGATCAAGAGGAGGAAAACCAAGAGGAGGAGGTCCACTCTTTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluGluGluGlnGluGlnGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGCTTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626

RESULT 3
ID ABU52412
XX ABU52412 standard; protein; 626 AA.
AC ABU52412;
XX
XX 10-MAR-2003 (first entry)
XX Peanut Ara h1 protein from cDNA clone p41b.
XX

KW Peanut; allergy; Ara h1; Ara h2; Ara h3; IgE binding site;
KW anaphylactic food allergen; anti-allergenic; vaccine; wound healing.
XX
OS Arachis hypogaea.
XX
PN WO200274250-A2.
XX
PD 26-SEP-2002.
XX
PF 18-MAR-2002; 2002WO-US009108.
XX
PR 16-MAR-2001; 2001US-0276822P.
PR 18-MAR-2002; 2002US-00276822.
XX
PA (PANA-) PANACEA PHARM.
XX
PI Caplan M, Sosin H, Sampson H, Bannan GA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabjohn PA, Shin DS, Stanley JS;
XX
XX WPI; 2003-018765/01.
XX N-PSDB; ABX70603.
XX
PT New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
XX
XX Claim 27; Fig 13; 300pp; English.
XX
CC The invention relates to a modified anaphylactic food allergen has an
CC amino acid sequence that is substantially identical to that of natural
CC anaphylactic food allergen, except for a cysteine residue that has been
CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IgE binding sites to
CC reduce allergenicity. Also included are: (1) a method of making a
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified
CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a
CC peanut allergen (e.g. Ara h1, h2 or h3)
XX
SQ Sequence 626 AA;

Alignment Scores:
Pred. No.: 1.59e-304 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.6% Indels: 0
DB: 6 Gaps: 0
US-10-728-323-1 (1-2032) x ABU52412 (1-626)

QY 50 ATGAGAGGGAGGGTTCCTCCACTGATGCTGTCTAGGATCCTGTCTGGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGCATGCCAAGTCATCACCTTACCAGAGAAACAGAGAACCCCTCGCCACG 169
Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGGTGCTCCAGAGTTGTCAACAGGAAACCGGATGACTTGAAGCAAAAGGCATCGAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGCTGCACCAAGCTCCAGTATGATCCTCGTTGTCTATGATCCTCGAGGACACACTGGC 289

Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAACGTTCCCTCCAGGGGACGACACAGTCGGCCGCAACCCGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCGTCAACCCGAAAGAGAGGAAGGCGCATGGGACCAGCTGGACCG 409
Db 101 AspAspArgArgGlnProArgGluGluGlyArgTrpGlyProAlaGlyPro 120
QY 410 AGGAGCGTCAAGAGAGAGACTGGAGACACACCAAGAGAACTTGGAGGCGACCAAGT 469
Db 121 ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgProSer 140
QY 470 CATCAGCAGCCACGGAATAAGCCCGAAGGAGAGAGAGAAACAAGAGTGGGAAACA 529
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGlnGluTrpGlyThr 160
QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCTTTTACTTCCCGTCAAGG 589
Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCGCTACGGGAACCAAAACGGTAGGATCCGGTCTCTGCAGAGGTTTGAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGGTCAGGCGAGTTTCAGAATCTCCAGAATCCAGTATTCGAGATCCGAGGCCAAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTTCTTCCCAAGCACGCTGTGATAACATCCTTGTATTCCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGCAGACCACTGCTAGCGTAGCAATGGCAATAACAGAAAGCTTTAATCTTGACGAG 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCATGCACTCAGAAATCCCATCCGGTTTCATTTCTTACATCTTGAACCGCCATGACAAC 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACTCAGAGTAGTAAATCTCATGCCCGTTAAACACACCCGCCCGCTTTGAGGAT 949
Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCAGCAGCGAGACCAATCATCTACTTCAGGCGCTTCAGCGAATACG 1009
Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCGCCTTCAATCGGAATTCATGAGATACGGAGGCTGCTGTAGAAGAGAT 1069
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGGAGAGGCGAGCGGATGGAGTACTCGAGTAGTGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGAGTATAGTCAAAGTGTCAAAGGACGCTTGAAAGAACTTACTAAGCAGCT 1189
Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCGCTTCAAGAAAGCTCCGAGAGAGGGAGATATCACCAACCCNATCAACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAGCGGAGCCCATCTTTCTAAACATTTGGGAAGTTATTTCAGGTGAAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGACCCCGCTTACGACCTGGACATGCTCACCCTGTGTAGAGATCAAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu 440

QY 1370 GGAGCTTTTGATGCTCCACACTTCAACTCAAGGCCATGGTTTATCTGCTCGTCAACAAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
QY 1430 GGAACCTGGAAAACCTTGAACTCGTGTGTAAGAAAAGAGCAACAACAGAGGGGACCGCGG 1489
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAGAGGAGGACGAGACCAAGAGAGGAGGAGTAACAGAGAGGTGCGTAGGTAC 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTTCAAGGAAGCGCATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGGCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGTTCTCCGAATCTCATCTGCTTGGCTTCGGTATCAACGCTGAAAAACCAACACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTGCAGGTGATAGGACAACTGTGATAGACCACATAGAGAACCAAGCGAAGTCT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTCCGGTCAAACAAGTTGAGAAAGCTCATCAAAAAACCAAGGAATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTCTCGTCTCTCAATCTCAATCTCAATCTCCGTCTCTGAGAAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTCGAAAAGAGATCAAGAGGAGGAAAACCAAGGAGGAGGTCCTCCTCTTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTTAAC 1927
Db 621 IleLeuLysAlaPheAsn 626
RESULT 4
ADC34906
ID ADC34906 standard; protein; 626 AA.
XX ADC34906;
AC ADC34906;
DT 18-DEC-2003 (first entry)
XX
DE Peanut allergen Ara h 1.
XX allergen; antigen; hyporesponsive; desensitisation; immunomodulator;
KW gene therapy; peanut.
XX
OS Arachis hypogaea.
XX
XX
PN WO2003047618-A2.
XX
PD 12-JUN-2003.
XX
XX 05-DEC-2002; 2002WO-GB005548.
XX
PR 05-DEC-2001; 2001US-0338385P.
XX
PA (CIRC-) CIRCASSIA LTD.
XX
PI Larche M, Ledger PW;
XX
DR WPI; 2003-523267/49.
XX
PT Desensitizing an individual to a selected polypeptide antigen comprises
PT administering a composition containing polypeptide antigens in an amount
PT that generates a state of hyporesponsiveness to the antigen to allow

desensitization.

Disclosure; Page 40; 57pp; English.

The invention relates to a novel method for desensitising an individual to a selected polypeptide antigen. The method comprises administering a composition that contains polypeptide antigens in an amount that generates in the individual a state of hyporesponsiveness to the antigen to allow desensitisation to one or more polypeptide antigens. The method of the invention has immunomodulator activity, and may have a use in gene therapy. The composition and method are useful in manufacturing a medicament for desensitising an individual to a selected polypeptide antigen or for generating in the individual a state of hyporesponsiveness to the antigen to allow desensitisation to one or more polypeptide antigens. The present sequence is used in the exemplification of the invention.

US-10-728-323-1 (1-2032) x ADC34906 (1-626)

Qy	50	ATGAGAGGAGGGTTTCTCCACTGATGCTGTGTCTAGGATCCCTTGTCTCTGCTTTCAGTT	109
Db	1	MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal	20
Qy	110	TCTCAACGCATGCCAAGTCATCACCTTACCAGAGAGAAAACAGAGAAACCCCTGCGCCACAG	169
Db	21	SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln	40
Qy	170	AGGTGCCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCT	229
Db	41	ArgCysLeuGlnSerCysGlnGlnLulProAspAspLeuLysGlnLysAlaCysGluSer	60
Qy	230	CGCTGCACCAAGCTCGAGTATGATCCTCGTGTGTCATGATCTCTCGAGGACACACTCGC	289
Db	61	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly	80
Qy	290	ACCACCAACCAACGTTTCCCTCCAGGGAGGGACACGTCGGCGCGCAACCCCGAGACTAC	349
Db	81	ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	100
Qy	350	GATGATGACCGCGCTCAACCCCGAAGAGAGAGAGCGCCGATGGGGACCACTGGACCG	409
Db	101	AspAspAspArgGlnProArgGlnProArgGluGluGlyArgTyrGlyProAlaGlyPro	120
Qy	410	AGGGAGCGTGAAGAGAGAGAGACTGGAGACAAACCAAGAGAAGATTGGAGGCGACCAAGT	469
Db	121	ArgGluArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer	140
Qy	470	CATCAGACGCCACGGAAATTAAGGCCCGAAGAGAGAGAGAGAAACAAGATGGGGAAACA	529
Db	141	HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr	160
Qy	530	CCAGGTAGCCATGTCAGGGAAGAAACATCTCGGAACAACCCCTTCTACTTCCCGTCAAGG	589
Db	161	ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg	180
Qy	590	CGGTTTAGCACCGCTACGGGAACCAAAACGCTAGGATCCGGGTCTCTGACAGAGTTTGAC	649
Db	181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	200
Qy	650	CAAGGTCAAGCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATTCAGGCCAAA	709
Db	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	220

Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCCTGACAAGAGGATCAAGAGGAGGAACCAAGAGGAGGAGGTCCTCTTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlySerGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTTAAC 1927
Db 621 IleLeuLysAlaPheAsn 626

RESULT 5
ADV97601
ID ADV97601 standard; protein; 626 AA.
XX
AC ADV97601;
XX
DT 10-MAR-2005 (first entry)
XX
DE Peanut Ara h I allergen alpha P17 seed storage protein.
XX
KW allergen; atopic dermatitis; diagnostic; immunotherapy;
KW immediate type hypersensitivity; Ara h I; dermatological; antiallergic;
KW immunosuppressive; antibody engineering; seed storage protein.
XX
OS Arachis hypogaea.
XX
PN US6835824-B1.
XX
PD 28-DEC-2004.
XX
PF 13-NOV-1998; 98US-00191593.
XX
PR 29-DEC-1995; 95US-0009455P.
PR 23-SEP-1996; 96US-00717933.
XX
PA (UYAR-) UNIV ARKANSAS.
XX
PI Burks AW, Stanley JS, Bannon GA, Cockrell G, Helm RM;
XX
WPI: 2005-045982/05.
DR N-PSDB; ADV97537, ADV97600.

PT New isolated nucleic acid encoding peanut allergen Ara h 1, useful for
PT obtaining diagnostics with optimal concentrations of each allergen, for
PT developing panels of mixtures of recombinant allergens, and in
PT immunotherapy.

PS Disclosure; Fig 16; 141pp; English.

CC This invention relates to a novel isolated nucleic acid molecule encoding
CC the peanut allergen Ara h I that has been identified in patients with
CC atopic dermatitis. Specifically, it refers to a monoclonal antibody
CC enzyme-linked immunosorbent assay (ELISA) used to detect and quantify the
CC peanut allergens, where the monoclonal antibodies have specificity for a
CC selected peanut allergen antigen. The present invention also describes
CC hybridomas that produce two-site monoclonal antibodies specific for a
CC peanut allergen that can be used in an ELISA to detect and determine the
CC concentration of a specific peanut allergen in a food product or in food
CC processing. Furthermore, it provides an expression vector and the
CC isolated nucleic acid or its fraction can include a diagnostic label.
CC Accordingly, these nucleic acids are useful for obtaining diagnostics
CC with optimal concentrations of each allergen or for developing panels of
CC mixtures of large numbers of recombinant allergens such that it can be
CC used in immunotherapy for the treatment of food hypersensitivity
CC reactions. This polypeptide sequence is the peanut Ara h I allergen
CC protein (P17 clone) of the invention.

XX Sequence 626 AA;

Alignment Scores: 1.59e-304 Length: 626
Pred. No.: 3286.00 Matches:
Score:

Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.6% Indels: 0
DB: 9 Gaps: 0
US-10-728-323-1 (1-2032) x ADV97601 (1-626)
QY 50 ATGAGAGGAGGTTTCTCCACTGATGCTGTGTGCTAGGATCCTTGTCTCGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGCATGCCAAGTCATCCTTACGAGAAGAAAACAGAGAACCCCTGCGCCAG 169
Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGGTGCTCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAACGAAAAGCATGCCAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGC 289
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAACGTTCCCTCCAGGAGCGGACACGTGGCCGCCAACCCGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCCGCTCAACCCCGAGAGAGAGGAAGGAGGCGCATGGTGGGACCACTGACCG 409
Db 101 AspAspAspArgArgGlnProArgGluGluGlyArgGlyArgGlnProGlyProAlaGlyPro 120
QY 410 AGGAGCGTGAAGAGAAGAGACTGGAGACAAACCAAGAGAAGATTGGAGGCGACCAAGT 469
Db 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATCAGCAGCCACGGAAAAATAAGCCCGAAGGAAGAGAGAGAACAGAGTGGGGAACA 529
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCCCTTCTACTTCCCGTCAAGG 589
Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCCGCTAGCGGAACCAAAACCGTAGGATCCGGGTCTCTCGAGAGTTTGAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGTCAAGCAGTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTGTGTTTCCCAAGCAGCTGATGCTGATAACATCCTTGTATTCCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTGACCGTAGCAATCGCAATAGCAATAACAGAAAGAGCTTTAATCTT 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCCATGACTCAGAAATCCCATCCCGTTTCATTTCTTCTATCATCTTGAACCGCCATGACA 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGCTAAAATCTCCATGCCGTTTAAACACACCCCGCCAGTTTGGAGAT 949
Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGGCGAGCAGCGGAGACCAATCATCTACTTTCGAGGGCTTCAGCAGGAATACG 1009
Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGAGGCGCGCTTCAATCGGAATTCAATGAGATACGAGGGTCTCTGTAGAAAGAAAT 1069

Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleAArgValLeuLeuGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGGAGAGGCGGATGGAGTACTCGGAGTAGTGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluAArgGlyGlnAArgTTPSerThrArgSerGluAsn 360
QY 1130 AATGAAGAGTAGTAGTCAAAAGTGTCAAGGAGCAGCTTGAGAACTTACTAAGCAGCT 1189
Db 361 AsnGluGlyValIleValIysValSerLysGluHisValGluLeuThrLysHisAla 380
QY 1190 AAATCCGTCTCAAGAAAGGCTCCGAGAGAGGAGAGATATACCAACCACTCAACTTG 1249
Db 381 LysSerValSerLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAGAAAGGCGAGCCGATCTTTCTAAACAATTGGGAAGTTATTGAGGTGAAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAACCCAGCTTCAGGACCTGGACATGCTACCTGTGTAGAGATCAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCACTCAAGGCCATGTTATCGTCGTCGCAACAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAACCTGGAAACCTTGAACCTCGTGTGAAGAAAGAGCAACAGAGGGGACGGCGG 1489
Db 461 GlyThrGlyAsnLeuGluValAlaValAArgLysGluGlnGlnArgGlyAArg 480
QY 1490 GAAGAGAGGAGCAGAGCAGAGAGAGAGGAGGAGTAACAGAGGTGCTAGGTAC 1549
Db 481 GluGluGluAspGluAspGluGluGluGlySerAsnArgGluValAArgGlyr 500
QY 1550 ACAGCGAGTTGAAGGAGCGATGTGTTTCATCATGCCAGCTCATCCAGTAGGCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleWetProAlaHisProValAlaIle 520
QY 1610 AAGCTTTCCTCCGAACCTCCATCTGCTTGGCTTCGTTATCAACGCTGAAACACACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTCTCTCAGGTGATAGGACAATGTCATAGACCATAGACAGAGCGAAGGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTGGGTCGGGTGAACAGTTGAGAAAGCTCATCAAAACACAGAGGAACT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTCTGAGTGCTGCTCTCAATCTCAATCTCAATCTCGGCTCTCTGAGAAAGAG 1849
Db 581 HisPheValSerAlaAArgProGlnSerGlnSerGlnSerProSerProGluLysGlu 600
QY 1850 TCTCTCTCAGAAAGAGGATCAGAGGAGGAGAAACCAAGGAGGAGGTCCACTCTTCTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626

RESULT 6

AAU04706
ID AAU04706 standard; protein; 626 AA.
XX AC
XX AAU04706;
XX AC
DT 23-OCT-2001 (first entry)
XX Anaphylactic antigen Ara h 1.
XX Ara h 1; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;
KW

allergy; mast cell; basophil; mouse.
Mus sp.
Key Location/Qualifiers
FT Misc-difference 285 /note= "Encoded by gta"
XX WO200140264-A2.
XX 07-JUN-2001.
XX 06-DEC-2000; 2000WO-US033124.
XX 06-DEC-1999; 99US-00455294.
XX 23-JUN-2000; 2000US-0213765P.
XX 27-SEP-2000; 2000US-0235797P.
XX (PANA-) PANACEA PHARM LLC.
XX (UYAR-) UNIV ARKANSAS.
XX (MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.
XX Bannan GA, Burks WA, Caplan MJ, Sampson H, Sosin H;
XX WPI; 2001-381378/40.
XX N-PSDB; AAS08537.
XX Antigenic fragments useful for reducing anaphylactic risk and reducing the severity and/or number of allergic symptoms in individuals sensitive to antigens, have reduced ability to bind Immunoglobulin E.
XX Claim 7; Fig 9; 100pp; English.
XX The sequence represents the amino acid sequence of anaphylactic antigen Ara h 1. Ara h 1 is an anaphylactic antigen (A), which was used to design antigenic peptides having a reduced ability to bind immunoglobulin E (IgE) as compared with the intact (A), or having a sequence substantially identical to a portion of sequence of an antigen that includes at least one IgE binding site, where at least one IgE binding site of the peptide is altered. The antigenic peptides are used in a composition which is useful for reducing risk or severity of allergic reaction to an antigen. This is done by identifying an individual at risk of allergic reaction to an antigen by identifying prior display of allergic symptoms when exposed to the antigen, or a familial relationship with an individual who previously displayed allergic symptoms when exposed to the antigen. Following this an antigen-specific IgE present on one or more mast cells or basophils in the individual's serum is identified. The individual is then contacted with a peptide corresponding to a portion of the antigen, which is selected, formulated, and delivered so that binding of the peptide to antigen-specific IgE is reduced as compared with IgE binding of intact antigen. The composition is also useful for treating and preventing allergic reactions
SQ Sequence 626 AA;
Alignment Scores: 3.84e-304 Length: 626
Pred. No.: 3282.00 Matches: 625
Score: 99.8% Conservative: 0
Percent Similarity: 99.8% Mismatches: 1
Best Local Similarity: 99.8% Indels: 0
Query Match: 91.5% Gaps: 0
DB: 4
US-10-728-323-1 (1-2032) x AAU04706 (1-626)
QY 50 ATCAGAGGGAGGTTTTCACACTGATGCTGTTCTAGGATCCTTCTCGGTTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGCATGCCAAGTCATCCCTTACCAGAAACACAGAACCCCTCGGCCAG 169
Db 21 SerAlaThrHisAlaLysSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40

QY 170 AGTGCCTCCAGAGTTGTTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATCGAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGTGCACCAAGCTCAGATATGATCTCGTTGTGTTCTATGATCCTCGAGGACACACTGGC 289
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAAGTTCCTCCAGGGGAGCGGACACACGTGGCGCCCAACCCGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATACCGCGTCAACCCGAGAGAGAGAGCGCGATGGGACACAGCTGACCG 409
Db 101 AspAspArgArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro 120
QY 410 AGGGAGCGTGAAGAGAAAGACTGGAGACAACCAAGAGAGATGGAGGGGACCAAGT 469
Db 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATCAGCAGCCGGAATAAGGCCGGAAGGAGAGAGAACCAAGAGTGGGGAACA 529
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGAACAACCCCTTCTACTTCCCGTCAAGG 589
Db 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCGCTACGGGAACCAAAACGGTAGGATCCGGGTCTCGAGAGTTTGAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGTCAAGGCAGTTTCAGAAATCTCCAGATCTCCAGATCACCGTATTGTGCAGATCGAGGCCAAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTTGTTCTCCAGCACGCTGTGATACATCTCTTATCCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTAGCCTAGCAATGGCAATACAGAAAGAGCTTTAATCTTGACGAG 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCATGCACTCAGAAATCCCATGCCGTTTCATCTCATCTTGAACCGCCATGACAAC 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACTCAGAGTAGCTAAATCTCCATGCCCGTTTAAACACACCCCGCCAGTTTGAGGAT 949
Db 281 GlnAsnLeuArgAlaAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCGGCGAGCGAGACCAATCATCTACTTCAGGGCTTCAGAGGAATACG 1009
Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCGCCCTTCAATCGGAATTCATAGATACGAGGCTGCTGTAGAGAGAAAT 1069
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGGTGACGAGAGAGAGAGGCGGAGCGGATGGAGTACTCGGAGTAGTGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTipSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGGAGTAGTAGTCAAGTGTCAAAGGACACGTTTGAAGAACTTACTAAGCAGCT 1189
Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCGCTCTCAAGAAAGGCTCCGGAAGAGAGGAGATATCACCAACCAATCAACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAGAAAGCGAGCCCGATCTTCTTAAACAACCTTTGGGAAGTTATTGAGGTGAAGCCAGAC 1309

Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAGAAGACCCACGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluLeuLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAAGGCGCATGGTTATCGTCTCGTCAACAAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAACCTGGAACCTTGAACCTCGTGTGTAAAGAGCAACAACAGAGGGGACGCGG 1489
Db 461 GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAAGAGGAGGACGAAGACGAAGAGGAGGAAGTAACAGAGAGAGGTGCGTAGCTAC 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTGAAGGAAGGCGATGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAACCTCATCTCGTTGGCTTCGGTATCAACGCTGAAAAACAACACACA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTCGAGGTGATAAGGACAAATGTGTAGACCAGATAGAGAAGCAAGCAAGGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTTCCCTGGTCCGGTCAACAAGTTGAGAAAGCTCATCAAAAACAGAGGAATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTAGTCTCGTCTCAATCTCAATCTCAATCTCCTCTCTCTCTCTCTCTCTCTCT 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTCTGAGAAGAGGATCAAGAGGAGGAACCAAGGAGGAAGGCTCCACTCTCTTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyGlyGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626
RESULT 7
ABU52570
ID ABU52570 standard; protein; 626 AA.
XX AC ABU52570;
XX DT 10-MAR-2003 (first entry)
XX DE Peanut Ara h1 mutant V72A.
XX KW Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;
KW muten; anaphylactic food allergen; anti-allergenic; vaccine;
XX wound healing.
XX OS Homo sapiens.
XX PN WO200274250-A2.
XX PD 26-SEP-2002.
XX PF 18-MAR-2002; 2002WO-US009108.
XX PR 16-MAR-2001; 2001US-0276822P.
XX PR 18-MAR-2002; 2002US-00276822.
XX PA (PANA-) PANACEA PHARM.

XX Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabjohn PA, Shin DS, Stanley JS;
XX
XX WPI; 2003-018765/01.
XX
XX New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
XX
XX Example 5; Page; 300pp; English.
XX
XX The invention relates to a modified anaphylactic food allergen has an
CC amino acid sequence that is substantially identical to that of natural
CC anaphylactic food allergen, except for a cysteine residue that has been
CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IgE binding sites to
CC reduce allergenicity. Also included are: (1) a method of making a
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified
CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a food
CC allergen, mutated to alter its IgE binding characteristics. Note: The
CC present sequence is not shown in the specification but was created by the
CC indexer using information provided in the specification
XX
SQ Sequence 626 AA;

Alignment Scores:

Pred. No.: 3,84e-304 Length: 626
Score: 3282.00 Matches: 625
Percent Similarity: 99.8% Conservativeness: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 91.5% Indels: 0
DB: 6 Gaps: 0

US-10-728-323-1 (1-2032) x ABUS2570 (1-626)

QY 50 ATGAGAGGAGGTTCTCCACTGATCTCTTCTAGGATCTCTGTCCTGCTTCTAGTT 109
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTGCAACGATGCCAAGTATCATCCTTACAGAGAAACAGAGAACCCCTGGCCAG 169
DB 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGGTGCTCCAGAGTCTGCAACAGAACCCGATGACTTGAAGAAAAGGCATGCGAGTCT 229
DB 41 ArgCysLeuGlnSerCysGlnGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGCTGCACCAAGCTCGAGTATGATCTCTGTTGTCTATGATCTCTGAGACACACTGGC 289
DB 61 ArgCysThrLysLeuGluTyrAspProArgCysAlaTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAAGCTTCCCTCCAGGGAGGGAGACAGTGGCCGCCCAACCCGGAGCTAC 349
DB 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATCATGACCGCTCAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 409
DB 101 AspAspAspArgArgGlnProArgArgGluGluGlyArgTrpGlyProAlaGlyPro 120
QY 410 AGGAGCGGTGAAG 469
DB 121 ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140

QY 470 CATCAGCAGCCACGGAATAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGluGluGluGluGluGlu 160
QY 530 CCAGGTAGCCATGTGAGGAG 589
DB 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCGCTACGGGAACCAAAACCGTAGAGTCCGGTCTCTGAGAGGTTTAC 649
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGTCAAGCAGTTTCCAGAAATCTCCAGAAATCACCGTATTGTGAGATCCAGGCCAAA 709
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTTGTCTTCTCCCAAGCAGCTGATGCTGATAACATCTCTTATCTCAGCAA 769
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTACCGTAGCAATGCAATACAGAAAGAGCTTTAATCTTACGAG 829
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCATGCACTCAGAAATCCATCCGTTTTCATTTCTACATCTTGAACCCCATGACAAC 889
DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACTCAGAGTAGTAAATCTCCATGCCCTTAAACACACCCCGCAGTTTGGAGAT 949
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTTCTCCCGCAGCAGCCGAGACCAATCATCTACTTCCAGGGCTTCCAGGAGATACG 1009
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCGCCCTTCAATCGGAATTCATAGATACGAGGAGTGTCTGTAGAGAGAAT 1069
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340
QY 1070 GCAGGAGGTGAGCAAG 1129
DB 341 AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAGGAGTGTAGTCAAGTGTCAAAGGAGCAGTTGAAGAACTTACTAGACAGCT 1189
DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluLeuLeuThrLysHisAla 380
QY 1190 AATCCGTCTCAAAGAAAGGCTCCGAAGAGAGAGAGAGATATCAACCAACCAACTTC 1249
DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAGCGAGCCCGATCTTTCTAACTTTGGGAAGTATTATGAGGTGAAGCCAGAC 1309
DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAACCCCGAGCTCAGGACCTGACATCTCTACCTGCTGTAGAGATCAAGAA 1369
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAAGCTTGTATGTCCTCCACACTTCAACTCAAGGCCATGGTTATCTGCTGCTCAACAA 1429
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAAGTGGAACTTGAAGTCTGCTGCTGTAAAGAAAAGAGCAACAGAGGGGACCGCG 1489
DB 461 GlyThrGlyAsnLeuGluValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAG 1549
DB 481 GluGluGluGluAspGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 500
QY 1550 ACAGCGAGGTGAAGGAAGGCGATGTTCATCATCGCAGAGCTCATCCAGTAGCCATC 1609

||||| 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
Db
||||| 1610 AAGCGTCTCCCAACTCCATCTGCTGGCTTCGGTATCAACGCTCAAAACACACACAGA 1669
Qy
||||| 521 AsnAlaSerSerGlnLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArg 540
Db
||||| 1670 ATCTTCCTTCGAGGTGATAGGCAATGTGATAGACCAGATAGAGCAAGCAAGGAT 1729
Qy
||||| 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
Db
||||| 1730 TTAGCATTCCTCGGTGCGGTGAACAAGTTGACAGCTCATCAAAACCAAGAGGATCT 1789
Qy
||||| 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
Db
||||| 1790 CACTTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCCGTCGTCCTCTGAGAAGAG 1849
Qy
||||| 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
Db
||||| 1850 TCTCTCGAGAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
Qy
||||| 601 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyLysGlyProLeuLeuSer 620
Db
Qy 1910 ATTTGAGGCTTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626
RESULT 8
ABU52568
ID ABU52568 standard; protein; 626 AA.
XX AC ABU52568;
XX DT 10-MAR-2003 (first entry)
XX DE Peanut Ara h1 mutant K32A.
XX KW Allergy; Ara h1; Ara h2; Ara h3; IGE binding site; peanut; mutant;
KW mutein; anaphylactic food allergen; antiallergenic; vaccine;
KW wound healing.
XX OS Homo sapiens.
XX PN WO200274250-A2.
XX PD 26-SEP-2002.
XX PF 18-MAR-2002; 2002WO-US0009108.
XX PR 16-MAR-2001; 2001US-0276822P.
XX PR 18-MAR-2002; 2002US-00276822.
XX PA (PANA-) PANACEA PHARM.
XX PI Caplan M, Sosin H, Sampson H, Bannon GH, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabinjohn PA, Shin DS, Stanley JS;
XX WPI; 2003-018765/01.
XX
XX New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
PT
XX Example 5; Page; 300pp; English.
XX
XX The invention relates to a modified anaphylactic food allergen has an
CC amino acid sequence that is substantially identical to that of natural
CC anaphylactic food allergen, except for a cysteine residue that has been
CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IGE binding sites to
CC reduce allergenicity. Also included are: (1) a method of making a
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food

CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified
CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a food
CC allergen, mutated to alter its IGE binding characteristics. Note: The
CC present sequence is not shown in the specification but was created by the
CC indexer using information provided in the specification
XX
SQ Sequence 626 AA;
Alignment Scores:
Pred. No.: 5,96e-304 Length: 626
Score: 3280.00 Matches: 625
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 91.5% Indels: 0
DB: 6 Gaps: 0
US-10-728-323-1 (1-2032) x ABU52568 (1-626)
Qy 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTTGCTAGGATCCTTGTCTCTGGCTTCAGTT 109
Db 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
Qy 110 TCTGCAACGCATGCCAAGTCATCCTTACCAAGAGAAACACAGAGAACCCCTCGGCCACG 169
Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnAlaLysThrGluAsnProCysAlaGln 40
Qy 170 AGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTCAAGCAAAAGCATCGAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60
Qy 230 CGCTGCACCAAGCTCGAGTATGATCCTCTGTTGTTCTATGATCCTCGAGACACACTGGC 289
Db 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
Qy 290 ACCCAACCAACGTTCCCTCCAGGGAGCGGACACGCTGGCGGCCCAACCCGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
Qy 350 GATGATGACCGCGCTCAACCCGAGAGAGAGGAGGCGCGATGGGACACAGCTGACCG 409
Db 101 AspAspAspArgArgGlnProArgArgGluGluGlyArgTrpGlyProAlaGlyPro 120
Qy 410 AGGGAGCGTGAAGAGAGAGAGACTGGAGACAACCAAGAGAGAGATTGGAGGCGCAAGT 469
Db 121 ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
Qy 470 CATAGCAGCCACCGGAAAATAAGGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
Qy 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCCGAAACACCTTTCTACTTCCCGTCAAGG 589
Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
Qy 590 CGGTTTAGCACCGCTTACCGGAAACCAAAACCGGTAGGATCCGGTCTCTCGAGAGGTTGAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGluArgPheAsp 200
Qy 650 CAAAGTCAAGCAGTTTCAGAAATCTCCAGAAATCCCGTATTGTGCAGATCGAGGCCAAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
Qy 710 CCTAACACCTTGTCTTCTCCCAAGCACGCTGATGCTGATAACATCTTGTATTCCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240

QY 770 GGCACGCCACCTGACCGTAGCAATGCGCAATAACAGAAAGAGCTTTAATCTTGACGAG 829
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 GlyGlnAlaThrValThrValAlaLaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY GGCATGCTCACTCAGAAATCCCATCCGGTTTCATTTCTTACATCTTGAAACCGCCATGACAAC 889
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY CAGAACTCAGAGTAGTAAATCTCCATGCCCGTTAAACACACCGCGCCAGTTTGAGGAT 949
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
281 GlnAsnLeuArgValAlaIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY TTCTTCCCGCAGCAGCGACCAATCATCTTACTTGCAGGCTTCAGCAGGAATACG 1009
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 PhePheProAlaSerArgAspGlnSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY TTGAGGCCCTTCAATGCGGAATCAATGAGATACGGAGGCTGCTGTAGAGAGAAT 1069
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340
QY GCAGGAGTGAGCAAGAGGAGAGGCGCAGCGGATGGAGTACTGGAGTAGTGAGAAC 1129
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
341 AlaGlyGlyGluGlnGluIleArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY AATGAGGAGTGATAGTCAAAGTGTCAAAGGAGCAGCTTGAAGAACTTACTAAGCAGCT 1189
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
361 AsnGlyGlyValIleValIleValIleValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY AAATCCCTCTCAAGAAAGGCTCCGAAGAGAGGAGATATCAACCAACCCCAATCAACTTG 1249
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY AGAAGAGCGGAGCCGAGTCTTTCTAACAATTTGGGAAGTATTAGGTGAAGCCAGAC 1309
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
401 ArgGlyGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY AAGAAGAACCCAGCTTCAGGACCTCGGACATGCTCACTGCTGAGAGATCAAGAA 1369
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY GGAGCTTTGTGCTCCACACTTCAACTCAAAGGCCATGTTATCGTCGTCGTCACAAA 1429
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY GGAAGTGAACCTTGAACTCGCTGCTGTAAGAAAGACGACAAACAGAGGGGACGCGG 1489
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArg 480
QY GAAGAAGAGGAGCAGCAAGCAAGAAAGAGGAGGGAAGTAACAGAGAGGTGCGTAGGTAC 1549
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
481 GluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValAlaArgGlyr 500
QY ACAGCAGGTTGAAGGAAGCGATGTGTTTCATCATGCCAGCTCATCCAGTAGCCATC 1609
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaHisProValAlaIle 520
QY AAGCTTCTCCGAACTCCATCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1669
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY ATCTTCTTCAGGTGATAAGCAATGTGATAGACAGATAGAGAGCAAGCAAGGAT 1729
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY TTAGCATTCCTCGGCTGGGTGAACAAAGTTGAGAGGCTCATCAAAAAACCAAGAGGAATCT 1789
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY CACTTTGTAGTGTCTGCTCTCAATCTCAATCTCGTCTGCTCTGCTGCTGCTGCTGCTGCT 1849
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
581 HisPheValSerAlaArgProGlnSerGlnSerProSerProSerProGluLysGlu 600
QY TCTCCTCAGAAAGAGGATCAAGAGGAGGAAAAACCAAGGAGGAGGCTCCACTCTCTTCA 1909

DB 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGCTTTTAAC 1927
DB 621 IleLeuLysAlaPheAsn 626
RESULT 9
ABU52571
ID ABU52571 standard; protein; 626 AA.
XX
AC ABU52571;
XX
XX 10-MAR-2003 (first entry)
DT
XX Peanut Ara h1 mutant R91A.
DE
XX Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;
KW mutein; anaphylactic food allergen; antiallergenic; vaccine;
KW wound healing.
XX Homo sapiens.
OS
XX WO200274250-A2.
PN
XX 26-SEP-2002.
PD
XX 18-MAR-2002; 2002WO-US009108.
PF
XX 16-MAR-2001; 2001US-0276822P.
PR 18-MAR-2002; 2002US-00276822.
XX
XX (PANA-) PANACEA PHARM.
PA
XX Caplan M, Sosin H, Sampson H, Bannon G, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabjohn PA, Shin DS, Stanley JS;
XX
XX WPI; 2003-018765/01.
XX
PT New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
XX
XX Example 5; Page; 300pp; English.
XX
CC The invention relates to a modified anaphylactic food allergen has an
CC amino acid sequence that is substantially identical to that of natural
CC anaphylactic food allergen, except for a cysteine residue that has been
CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IgE binding sites to
CC reduce allergenicity. Also included are: (1) a method of making a
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified
CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a food
CC allergen, mutated to alter its IgE binding characteristics. Note: The
CC present sequence is not shown in the specification but was created by the
CC indexer using information provided in the specification
XX
SQ Sequence 626 AA;

Alignment Scores: 5.96e-304 Length: 626
Pred. No.: 3280.00 Matches: 625
Score: 99.8% Conservative: 0
Percent Similarity: 99.8% Mismatches: 1
Best Local Similarity: 99.8%

Query Match:	91.5%	Indels:	0
DB:	6	Gaps:	0
US-10-728-323-1 (1-2032) x ABUS2571 (1-626)			
QY	50	ATGAGAGGAGGTTTCTTCACGTATGCTGTTCTAGGATCCTTCTCTGGCTTCAGTT	109
DB	1	MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal	20
QY	110	TCGCAACGCATCCCAAGTATCATCCTTACCAGAAGAAACAGAGAACCCCTCGCCAG	169
DB	21	SerAlaThrHisAlaIlyssSerProTyrGlnLysLysThrGluAsnProCysAlaGln	40
QY	170	AGTGTGCTCCAGAGTTGTTCAACAGAACCCGATGACTTGAAGCAAAAGGCATCCGAGTCT	229
DB	41	ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer	60
QY	230	CGCTGCACCAAGCTCAGATATGATCTCTGTTGTCTATGATCTCCTCAGAGACACATGGC	289
DB	61	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly	80
QY	290	ACCACCAACCAAGTTCCCTCCAGGGGAGCGACACGTCGGCCCAACCCCGGAGACTAC	349
DB	81	ThrThrAsnGlnArgSerProGlyGluAlaThrArgGlyArgGlnProGlyAspTyr	100
QY	350	GATGATGACCCGCTCAACCCCGAAGAGAGGAAGGCGCGATGGGACCCAGCTGGACCG	409
DB	101	AspAspAspArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro	120
QY	410	AGGAGCGGTAAAGAGAAGAAGACTCGAGACAAACCAAGAGAAGATTGGAGGCGACCAAGT	469
DB	121	ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgProSer	140
QY	470	CATCACGCCCGGAAATAGCCCGAGGAGGAGAGAGAGAGAGAAACAAGAGTGGGAAACA	529
DB	141	HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGlnGlnGluTrpGlyThr	160
QY	530	CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCTTTCTACTTCCCGTCAAGG	589
DB	161	ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg	180
QY	590	CGGTTTAGCACCCGCTACGGGAACCAAAACGGTAGGATCCGGTCTCTGAGAGTTTGAC	649
DB	181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	200
QY	650	CAAGGCTCAGGCAGTTTCAGATCTCCAGATCCACGATTCGTCAGATCCGAGGCCAAA	709
DB	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys	220
QY	710	CCTAACACTCTTCTTCTCCAGCACGCTGATGCTGATAACATCCTTGTATCTCCAGCAA	769
DB	221	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	240
QY	770	GGCAAGCACCGTACCGTAGCAATGGCAATACAGAAAGAGCTTTAATCTTGACGAG	829
DB	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260
QY	830	GGCCATGCATCAGATCCCATCCGCTTCATTTCTTACATCTTGACCCGCTGACAC	889
DB	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280
QY	890	CAGAACCTCAGAGTAGTAAATCTCCATGCCCGTTAAACACACCCCGCCAGTTTGAGGAT	949
DB	281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	300
QY	950	TTCTTCCCGCGAGCAGCCAGACCAATCATCTTCTGAGGGCTTCAGCAGGAATACG	1009
DB	301	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	320
QY	1010	TTGGAGCCGCTTCAATCGGNAATTCAATGAGATACGAGGGTGTCTGTAGAGAGAT	1069
DB	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn	340

QY	1070	GCAGGAGGTGAGCAAGAGGAGAGAGCGGATCGAGTACTCGAGTAGTGAGAAC	1129
DB	341	AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn	360
QY	1130	AATGAGGAGTGATAGTCAAAAGTGTCAAAGAGACAGTGTGAAGAACTTACTAAGCAGCT	1189
DB	361	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380
QY	1190	AAATCCGCTCTCAAGAAAGGCTCCCAAGAGAGGAGATATCACCAACCCCAACTTCTG	1249
DB	381	LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu	400
QY	1250	AGAGAGCGCAGCCCATCTTTTAAACACTTTTGGGAAGTTATTTGAGGTGAAGCAGAC	1309
DB	401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420
QY	1310	AGAAGAACCCCGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA	1369
DB	421	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu	440
QY	1370	GGAGCTTTGATGCTCCCACTTCAACTCAAGGCCATCGTTTATCGTCTCGTCAACAA	1429
DB	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys	460
QY	1430	GGAACCTGGAACCTTTGAACTCGTGGCTGTAAAGAAAGACAAACAAGAGGGGACGCGG	1489
DB	461	GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnArgGlyArgArg	480
QY	1490	GAAGAAGGAGGAGCGAAGACCAAGAGGAGGAGGAGTAACAGAGAGGTGCGTAGGTAC	1549
DB	481	GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr	500
QY	1550	ACAGCGAGGTTCAGGAGGCGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGGCATC	1609
DB	501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaHisProValAlaIle	520
QY	1610	AACGCTTCTCCGAACTCCATCTGCTTGGCTTCGCTATCAACGCTGAAACCAACACAGA	1669
DB	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg	540
QY	1670	ATCTTCTTGCAGGTGATAAGACAAATGTGATAGACCATAGAGNAGCAAGCAAGGAT	1729
DB	541	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp	560
QY	1730	TTAGCATTCCTCGGTGCGGTGAACAAGTTGAGAAGCTCATCAAAAACAGAGGAATCT	1789
DB	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	580
QY	1790	CACCTTGTGAGTGCCTCCTCAATCTCAATCTCAATCTCCGCTCCTCGAGAAAGAG	1849
DB	581	HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu	600
QY	1850	TCTCTGAGAAAGAGATCAAGAGGAGGAAACCAAGGAGGAGGAGGTCCTCTCTTCA	1909
DB	601	SerProGluLysGluAspGlnGluGluAsnGlnGlyLysGlyProLeuLeuSer	620
QY	1910	ATTTTGAAGGCTTTTAAAC	1927
DB	621	IleLeuLysAlaPheAsn	626
RESULT 10			
ID	ABUS2574	standard; protein; 626 AA.	
XX	ABUS2574;		
XX	10-MAR-2003 (first entry)		
DE	Peanut Ara h1 mutant R499A.		
KW	Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;		
KW	mutain; anaphylactic food allergen; antiallergenic; vaccine;		
KW	wound healing.		

XX OS Homo sapiens.
XX PN WO200274250-A2.
XX PD 26-SEP-2002.
XX PF 18-MAR-2002; 2002WO-US009108.
XX PR 16-MAR-2001; 2001US-0276822P.
XX PR 18-MAR-2002; 2002US-00276822.
XX PA (PANA-) PANACEA PHARM.
XX PI Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
XX PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
XX PI Rabjohn PA, Shin DS, Stanley JS;
XX XX
XX WI; 2003-018765/01.
XX
XX New modified anaphylactic food allergen, useful for preventing or
XX treating allergic reactions associated with e.g. anaphylactic allergens.
XX
XX Example 5; Page; 300pp; English.
XX
XX The invention relates to a modified anaphylactic food allergen has an
XX amino acid sequence that is substantially identical to that of natural
XX anaphylactic food allergen, except for a cysteine residue that has been
XX modified so that it cannot participate in the disulphide bond. The
XX modification may also comprise mutation of the IgE binding sites to
XX reduce allergenicity. Also included are: (1) a method of making a
XX modified anaphylactic food allergen; (2) a nucleotide molecule encoding
XX or for causing a site specific mutation in the modified anaphylactic food
XX allergen; (3) a transgenic plant or animal expressing the modified
XX anaphylactic food allergen; (4) a method of treating an individual by
XX reducing the clinical response to a natural anaphylactic food allergen;
XX and an isolated fragment of peanut allergen Ara h 1. The modified
XX anaphylactic food allergen is useful for preventing or treating allergic
XX reactions associated with any natural allergen such as food, insect,
XX rubber or preferably anaphylactic allergens. It is also useful for
XX treating wounds in mammals such as bovine, canine, feline, caprine,
XX ovine, porcine, murine or equine species. The present sequence is a food
XX allergen, mutated to alter its IgE binding characteristics. Note: The
XX present sequence is not shown in the specification but was created by the
XX indexer using information provided in the specification
XX
XX Sequence 626 AA;

Alignment Scores:
Pred. No.: 5.96e-304 Length: 626
Score: 3280.00 Matches: 625
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 91.5% Indels: 0
DB: 6 Gaps: 0

US-10-728-323-1 (1-2032) x ABU52574 (1-626)
QY 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTGTCTAGGGATCTCTGTCTCGGTTTCAGTT 109
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyLeuValLeuAlaSerVal 20
QY 110 TGTGCAACGGATGCCAAGTCACTTACAGAGAAACAGAGAACCCCTCGGCCAG 169
DB 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGGTGCTCCAGAGTTGTCACAGAACCCGATGACTTGAAGCAAAAGGCATCGAGTCT 229
DB 41 ArgCysLeuGlnSerCysGlnGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGTGTCCACCAAGCTCGAGTATGATCCCTCGTTGTGTCTATGATCCTCGAGACACACTGGC 289
DB 61 ArgCysThrLysLeuGluTyrAspProArgProArgCysValTyrAspProArgGlyHisThrGly 80

QY 290 ACCACCAACCAAGCTTCCCTCCAGGGAGCGGACACACGTCGGCCCAACCCGGAGACTAC 349
DB 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCGTCAACCCCGAAGAGAGAGGAGCGCGATGGGAGCAGCTGACCG 409
DB 101 AspAspArgArgGlnProArgArgGluGluGlyGlyArgTyrGlyProAlaGlyPro 120
QY 410 AGGAGCGTGAAGAGAAAGACTGGAGACAAACCAAGAGAGATTTGGAGGCGACCAAGT 459
DB 121 ArgGluArgGluArgGluAspTyrArgGlnProArgGluAspTyrArgArgProSer 140
QY 470 CATCAGCAGCCCGGAAATAAGCCCGAAGAGAGAGAGAGAGAAACAGAGTGGGAGACA 529
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTyrGlyThr 160
QY 530 CCAGGTAGCATGTGAGGGAAGAAACATCTCGGAACAAACCTTTCTACTTCCCGTCAAGG 589
DB 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCGCTACCGGAAACCAAAACGGTAGGATCCGGGTCTCGAGAGTTTGAC 649
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGTCAAGCAGTTTCAGATCTCCAGATCACCGTATTCTGCAGATCGAGGCCAAA 709
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTGTTCTTCCCAAGCACGCTGTGCTGATAACATCTCTGTTATCCAGCAA 769
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTAGCCCTAGCAATGCGCAATAACAGAAAGAGCTTTAATCTTGACGAG 829
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCATGCACTCAGATCCATCCGCTTTCATTTCTTACATCTTGAACGCCCATGACAAC 889
DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACTCAGAGTAGCTAAAATCTCCATGCCCGTTAAACACACCCGCCAGTTTGAAGAT 949
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGAGCGAGACCAATCATCTACTTTCAGGCGCTTCAGCAGGAATACG 1009
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCGCCCTTCAATCGGNAATTCAATGAGATACGGAGGCTGCTGTTAGAGAGAAT 1069
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGAGAGAGCGGAGCGAGCTGAGTAGTACTCGGAGTAGTGAAC 1129
DB 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTyrSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGGAGTGTAGTCAAAAGTGTCAAAGGAGCAGCTTGAAGAACTTACTAGCAGCT 1189
DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCGTCTCAAGAAAGCTCCGAGAGAGAGGAGATATCACCAACCCCAATCAACTTG 1249
DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAGCGGAGCCCGATCTTTCTAAACACTTTGGGAAGTATTATTGAGGTGAAGCCAGAC 1309
DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTGCTGTAGAGATCAAGAA 1369
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440

QY 590 CGGTTTAGCACCCGCTACCGGAAACCAAAACGGTAGGATCCGGTCTCGCAGAGTTTGAC 649
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGGTCAAGGCAGTTTCAGAAATCCAGAAATCACCGTATTGTGCAGATCGAGGCCAAA 709
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAAACACTCTGTCTCTCCCAAGCAGCTGATGCTGATAACATCTCTGTATCCAGCAA 769
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTAGCCGAGCAAAATGGCAATAACAGAAAGAGCTTTATCTTGACGAG 829
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GSCCATGCATCAGAACTCCATCCGTTTTCATTTCTTACATCTTGAACGCCCATGACAAC 889
DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACTCAGAGTAGTAAATCTCCATGCCCGTTAAACACACCCGCCAGTTTGAGGAT 949
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGAGCAGCCGAGACCAATCATCTTACTTGAGCGCTTCAGCAGGAATACG 1009
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGGAGCCCGCTCAATCGGATTCATGAGATCAGGAGGCTGCTGTAGAGAGAT 1069
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGGTGAGCAAGAGGAGAGGCGATGGAGTACTCGAGTAGTGAGAAC 1129
DB 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGAGTAGTAGTCAAAAGTGTCAAAGAGCACCTTGAAGAACTTACTAGACGCT 1189
DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCCTCTCAAGAAGGCTCCGAGAAGAGGAGATATCACCAACCCCAACTCAGTTG 1249
DB 381 LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGCGCAGCCCGATCTTTCTTAAACAACCTTTGGGAAGTTATTGTAGGTGAAGCCAGAC 1309
DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAAGACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAAGGCCATGTTATCGTCGTCGCAACAAA 1429
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAACCTGGAACCTTGAACCTCGTGGCTGTAAAGAAAGAGCAACACAGAGGGGACCGCGG 1489
DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAAGAGGAGGAGCGAAGCAAGAGAGAGGAGGAAGTAAACAGAGGTGCGTAGGTAC 1549
DB 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValAlaArgTyr 500
QY 1550 ACAGGAGTTGAAGGAGGCGATGTGTTCATCATCGCAGCAGCTCATCAGTAGCCATC 1609
DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAACTCCATCTCTGCTTGGCTTCGGTATCAACGCTGAAAAACAACACAGA 1669
DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540

QY 1670 ATCTTCTTGCAGGTGATTAAGGACAAATGTGATAGACCAGATAGAGCAAGCGAGGAT 1729
DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCAATTCCTTGGGTCCGGTCAACAAAGTTGAGAAGCTCATCAAAAACAGAGGAATCT 1789
DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTCTGCTCCTCAATCTCAATCTCAATCTCCTGCTGCTCTGAGAAAGAG 1849
DB 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTCGAAGAGAGATCAAGGAGGAGAAACCAAGGAGGAGGAGGCTCCACTCCTTTCA 1909
DB 601 SerProGluLysGluAspGlnGluGluGlnGlnGlnGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTTAAAC 1927
DB 621 IleLeuLysAlaPheAsn 626
RESULT 12
ADM12053
ID ADM12053 standard; protein; 626 AA.
XX ADM12053;
XX
XX 20-MAY-2004 (first entry)
XX
XX Arachis hypogaea 1 (Ara h1) protein.
XX
XX antigen presentation enhancing hybrid polypeptide; mammalian II-Key;
XX MHC class II; antibacterial; virucide; fungicide; antirheumatic;
XX antiarthritic; neuroprotective; dermatological; immunosuppressive;
XX antiinflammatory; antidiabetic; antithyroid; immune;
XX rheumatoid arthritis; multiple sclerosis; lupus erythematosus;
XX diabetes mellitus; myasthenia gravis; autoimmune thyroiditis;
XX scleroderma; dermatomyositis; pemphigus.
XX
XX Arachis hypogaea.
XX
XX US2003235594-A1.
XX
XX 25-DEC-2003.
XX
XX 17-SEP-2002; 2002US-00245871.
XX
XX 14-SEP-1999; 99US-00396813.
XX
XX 17-JUL-2002; 2002US-00197000.
XX
XX (ANTI-) ANTIGEN EXPRESS INC.
XX
XX Humphreys R, Xu M;
XX
XX WPI; 2004-070554/07.
XX
XX Novel II-Key/antigen presentation enhancing hybrid polypeptide, useful
XX for treating infections, rheumatoid arthritis, multiple sclerosis, lupus
XX erythematosus and diabetes mellitus.
XX
XX Example 1; Page 16-17; 87pp; English.
XX
XX The invention relates to a novel antigen presentation enhancing hybrid
XX polypeptide. The novel polypeptide has an N-terminal element consisting
XX of 4-16 residues of a mammalian II-key peptide and its non-N-terminal
XX deletion modifications, a chemical structure covalently linking the N-
XX terminal element to an MHC class II-presented epitope of a C-terminal
XX element. The C-terminal element comprises an antigenic epitope, which
XX binds to an antigenic peptide binding site of an MHC class II molecule.
XX The antigen presentation enhancing hybrid polypeptide has the following
XX activities: antibacterial, virucide, fungicide, antirheumatic,
XX antiarthritic, neuroprotective, dermatological, immunosuppressive,
XX antiinflammatory, antidiabetic, and antithyroid. The antigen presentation
XX enhancing hybrid polypeptide is useful for modulating the immune response

CC in an individual and for treating infections (such as bacteria, virus,
 CC parasite and fungus), rheumatoid arthritis, multiple sclerosis, lupus
 CC erythematous, diabetes mellitus, myasthenia gravis, autoimmune
 CC thyroiditis, scleroderma, dermatomyositis and pemphigus. This sequence
 CC represents a mammalian II key related protein of the invention.

XX
 SQ Sequence 626 AA;

Alignment Scores:
 Pred. No.: 7, 43e-304 Length: 626
 Score: 3279.00 Matches: 625
 Percent Similarity: 99.8% Conservative: 0
 Best Local Similarity: 99.8% Mismatches: 1
 Query Match: 91.4% Indels: 0
 DB: 8 Gaps: 0

US-10-728-323-1 (1-2032) x ADM12053 (1-626)

QY 50 ATGAGAGGAGGTTCTCCACTGATGCTTCTAGGATGCTTGTCTGCTTTCAGTT 109
 DB 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyLeuValLeuAlaSerVal 20
 QY 110 TCTGCAACGATCCCAAGTATCACCTTACCAGAGAAACACAGAGACCCCTCGCCAG 169
 DB 21 SerAlaThrHisAlaLysSerProTyrGlnLysThrGluAsnProCysAlaGln 40
 QY 170 AGGTGCTCCAGAGTTGTCAACAGGAACCGATGACTTGAAGCAAAAGGCGATCCGAGTCT 229
 DB 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlyLysAlaCysGluSer 60
 QY 230 CGTGCACCAAGTCCAGTATGATCTCTGCTGTTGTCTATGATCCTCGAGGACACACTGGC 289
 DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
 QY 290 ACCACCAACCAAGTTCCTCCAGGAGCGGACAGCTGGCCGCCAACCCGAGACTAC 349
 DB 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
 QY 350 GATGATGACCGCGCTCAACCCGAGAGAGAGGAGCGCGATGGGACCGAGCTGGACCG 409
 DB 101 AspAspArgArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro 120
 QY 410 AGGAGCGTGAAGAGAGAAAGACTGGAGACAAACCAAGAGAGATTGGAGGCGACCAAGT 469
 DB 121 ArgGluArgGluArgGluArgTrpArgGlnProArgGluAspTrpArgProSer 140
 QY 470 CATCAGCAGCCAGGAATAAGCCCGAGGAGAGAGAGAGAGAAACAAGATGGGAAACA 529
 DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
 QY 530 CCAGGTAGCCATGTGAGGGAAGAAACATCTCGGAACAAACCTTCTACTTCCCGTCAAGG 589
 DB 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
 QY 590 CGGTTTGAACCCGCTACCGGAACCAAAACGGTAGGATCCGGTCTCTCGAGAGGTTTGAC 649
 DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
 QY 650 CAAGGTCAAGGCGAGTTTCAAGATCTCCAGATCCCGATTTGTGAGATCGAGGCCAAA 709
 DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
 QY 710 CCTAAACACTTGTCTTCCCAAGCAGCTGATGCTGATAACATCTTGTATTATCCAGCAA 769
 DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnLeuValIleGlnGln 240
 QY 770 GGGCAAGCCACCGTAGCCGTAAGTAATGGCAATTAACAGAGAAAGACTTAACTTTGACGAG 829
 DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
 QY 830 GGCATGCACTCAGATCCCATCCGGTTTCATTTCTTCTACATCTTGAACCCCATGACAAC 889
 DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280

QY 890 CAGAACCTCAGAGTAGCTAAAAATCTCCATGCCCGTTAAACACACCCGCGCAGTTTGGAGT 949
 DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
 QY 950 TTCTTCCCGCGAGCAGCCGAGACCAATCATCTACTTCAGCGGCTTCAGCAGGAATACG 1009
 DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
 QY 1010 TTGGAGGCGCGCTCAATCGCGAATTCAATGATAGATACGAGGGTCTGTAGAGAGAAT 1069
 DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
 QY 1070 GCAGGAGGTGAGCAAGAGAGAGAGCGGAGAGCCATGATCGAGTACTCGGAGTAGTGAGAAC 1129
 DB 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgTrpSerThrArgSerSerGluAsn 360
 QY 1130 AATGAAGGAGTATAGTCAAGTGTCAAGGAGACGCTTGAAGAACTTACTAAGCAGCT 1189
 DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
 QY 1190 AATCCGCTCAAGAAAGGCTCCGAGAGAGAGGAGATATCACCAACCCCACTCAACTTG 1249
 DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
 QY 1250 AGAAGAGCGAGCCCGATCTTCTAACTTTGGGAAGTATTTCAGGTGAAGCCAGAC 1309
 DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
 QY 1310 AAGAAGAACCCCGACTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
 DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
 QY 1370 CGAGCTTTGATGCTCCACACTTCAACTCAAGAGCCATCGTTATCGTCTCGTCAACAA 1429
 DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
 QY 1430 GGAATCGAAACTCTGCTGTAAAGAAAGACAAACAGAGGGGACCGCGG 1489
 DB 461 GlyThrGlyAsnLeuLeuValAlaValArgLysGluGlnGlnArgGlyArgGlyArg 480
 QY 1490 GAAGAAGGAGGAGCAAGACGAGAGAGAGAGAGGAGTAACAGAGAGGTGCGTAGGTAC 1549
 DB 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
 QY 1550 ACAGCGAGGTTCAGAGGCGCATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGGCATC 1609
 DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaHisProValAlaIle 520
 QY 1610 AACGCTTCTCCGAACTCCATCTGCTTGGCTTCGCTATCAACGCTGAAACCAACACAGA 1669
 DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
 QY 1670 ATCTCTCTTGCAGGTATTAAGACCAATGTGATAGACCATAGAGAGAGCAAGCAAGGAT 1729
 DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
 QY 1730 TTGATCTCCCTGGTCCGGTCAACAGTTGAGAGCTCATCAAAACCCAGAGGAATCT 1789
 DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
 QY 1790 CACTTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCCGCTCTCCCTGAGAAAGAG 1849
 DB 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
 QY 1850 TCTCTGAGAAAGAGATCAAGAGAGGAGAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
 DB 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
 QY 1910 ATTTTGAAGCTTTTAAAC 1927
 DB 621 IleLeuLysAlaPheAsn 626

RESULT 13
ABU52569
ID ABU52569 standard; protein; 626 AA.
XX
AC ABU52569;
XX
DT 10-MAR-2003 (first entry)
XX
DE Peanut Ara h1 mutant D52A.
XX
KW Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;
KW mutein; anaphylactic food allergen; antiallergenic; vaccine;
KW wound healing.
XX
OS Homo sapiens.
XX
PN W0200274250-A2.
XX
PD 26-SEP-2002.
XX
PF 18-MAR-2002; 2002WO-US009108.
XX
PR 16-MAR-2001; 2001US-0276822P.
PR 18-MAR-2002; 2002US-00276822.
XX
PA (PANA-) PANACEA PHARM.
XX
PI Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabinjohn PA, Shin DS, Stanley JS;
XX
XX WPI; 2003-018765/01.
XX
PT New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
XX
PS Example 5; Page; 300pp; English.
XX
CC The invention relates to a modified anaphylactic food allergen has an
CC amino acid sequence that is substantially identical to that of natural
CC anaphylactic food allergen, except for a cysteine residue that has been
CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IgE binding sites to
CC reduce allergenicity. Also included are: (1) a method of making a
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified
CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a food
CC allergen, mutated to alter its IgE binding characteristics. Note: The
CC present sequence is not shown in the specification but was created by the
CC indexer using information provided in the specification
XX
SQ Sequence 626 AA;

Alignment Scores:
Pred. No.: 9.26e-304 Length: 626
Score: 3278.00 Matches: 625
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 91.4% Indels: 0
DB: 6 Gaps: 0

US-10-728-323-1 (1-2032) x ABU52569 (1-626)

QY 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTGTAGGATCCCTTGTCTGGCTTCAGTT 109
|||||

1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal 20
110 TCTGCAACGCATGCCAAGTCATCCTTACCAAGAAACACAGAGAACCCCTCGCCAG 169
|||
21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
|||
170 AGTGCTCTCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGCATGCCAGTCT 229
|||
41 ArgCysLeuGlnSerCysGlnGlnGluProAspAlaLeuLysGlnLysAlaCysGluSer 60
|||
230 CGCTGCACCAAGCTCGAGTATGATCCTCTGTGTCTATGATCTCTCAGAGACACACTGGC 289
|||
61 ArgCysThrLysLeuGlnTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
|||
290 ACCACCAACCAAGTTCCTCCAGGGAGCGGACACGTCGGCGCCCAACCCGAGACTAC 349
|||
81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
|||
350 GATGATGACCGCGCTCAACCCGAGAGAGAGAGGAGGCGGATGGGACACAGCTGACCG 409
|||
101 AspAspArgArgGlnProArgArgGluGluGlyArgTrpGlyProAlaGlyPro 120
|||
410 AGGGAGCGTAAAGAGAGAGAGAGCTGGAGACAAACCAAGAGAGATTGGAGGCGACCAAGT 469
|||
121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgProSer 140
|||
470 CATCAGCAGCCCGGAAATAAGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
|||
141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
|||
530 CCAGTTAGCATGTGTAGGGAAGAAACATCTCGGAACAAACCTTTCTACTTCCCGTCAAGG 589
|||
161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
|||
590 CGGTTTGTAGCACCGCTACGGGAACCAAAACGGTAGGATCCGGGTCTCTGAGAGTTTGAC 649
|||
181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
|||
650 CAAGGTCAAGCAGGTTTCAGATCTCCAGATCACCGTATTCTGCAGATCCGAGGCCAAA 709
|||
201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
|||
710 CCTAACACATCTTCTCTCCAGACGCTGATGCTGATAACATCTCTGTTATCTCCAGCAA 769
|||
221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
|||
770 GGCAAGCCACCGTACCGTACCAATGGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 829
|||
241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
|||
830 GGCATGCACTCAGATCCATCCGCTTTCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 889
|||
261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
|||
890 CAGAACTCAGAGTAGTCTAAATCTCATGCGCCGTTAAACACACCCCGCCAGTTTGGAGAT 949
|||
281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
|||
950 TTCTTCCCGCGAGCGGAGACCAATCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1009
|||
301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
|||
1010 TTGAGGCGCCCTTCAATCGGAATTCATGATACGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1069
|||
321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
|||
1070 GCAGAGGAGTGCAGCAAGAGAGAGAGGCGGAGGCGGAGTGGAGTACTCGAGTAGTGAAC 1129
|||
341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn 360
|||
1130 AATGAGGAGTGTAGTCAAAAGTGTCAAGGAGCAGCTTCAAGAACTTACTTACGACGCT 1189
|||
361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
|||

QY 1190 AAATCCGTCTCAAAGAAAGGCTCCGAAGAAGAGGAGATATCACCAACCCAAATCAACTTG 1249
DB |||||||
DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAGGAGCCCGATCTTCTTAACAACATTTGGGAAGTTATTGAGGTGAAGCCAGAC 1309
DB |||||||
DB 401 ArgGluGluProAspLeuSerAsnAsnAsnGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAACCCCGAGCTTCAGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
DB |||||||
DB 421 LysLysAsnProGlnLeuAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATCTCCACACTTCAACTCAAGGCCATGTTATCTGCTCGTCAACAAA 1429
DB |||||||
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAATCGAAACCTTGAATCTGTGGCTGTAAGAAAAGACAAACAGAGGGGAGCGGG 1489
DB |||||||
DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAGAGGAGACCAAGACCAAGAGAGAGGAGGAGTAACAGAGAGGTGCGTAGGTAC 1549
DB |||||||
DB 481 GluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTGAAGGAAGCGGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGGCATC 1609
DB |||||||
DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGTTCTTCGAATCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1669
DB |||||||
DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTCAGGTATAGGACATCTGATAGACCATATAGACCATATAGACCATATAGACCAT 1729
DB |||||||
DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTGCGGTGAACAAAGTTGAGAAGCTCATCAAAAACAGAGGAATCT 1789
DB |||||||
DB 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTGCTCGTCTCAATCTCAATCTCAATCTCAATCTCGTCTCGTCTGAGAAAGAG 1849
DB |||||||
DB 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTGAGAAAGAGATCAAGAGGAGGAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
DB |||||||
DB 601 SerProGluLysGluAspGlnGluGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGCTTTTAAAC 1927
DB |||||||
DB 621 IleLeuLysAlaPheAsn 626

RESULT 14
ABU52572
ID ABU52572 standard; protein; 626 AA.
XX
XX AC ABU52572;
XX
XX 10-MAR-2003 (first entry)
XX Peanut Ara h1 mutant D103A.
XX
KW Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;
KW mutein; anaphylactic food allergen; anti-allergenic; vaccine;
KW wound healing.
XX
OS Homo sapiens.
XX
XX PN WO200274250-A2.
XX
XX 26-SEP-2002.
XX

PF 18-MAR-2002; 2002WO-US009108.
XX
PR 16-MAR-2001; 2001US-0276822P.
PR 18-MAR-2002; 2002US-00276822.
XX (PANA-) PANACEA PHARM.
XX Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabjohn PA, Shin DS, Stanley JS;
XX WPI; 2003-018765/01.
XX
PT New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
XX
XX Example 5; Page: 30pp; English.
XX
XX The invention relates to a modified anaphylactic food allergen has an
CC amino acid sequence that is substantially identical to that of natural
CC anaphylactic food allergen, except for a cysteine residue that has been
CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IgE binding sites to
CC reduce allergenicity. Also included are: (1) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified
CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a food
CC allergen, mutated to alter its IgE binding characteristics. Note: The
CC present sequence is not shown in the specification but was created by the
CC indexer using information provided in the specification
XX
SQ Sequence 626 AA;

Alignment Scores:
Pred. No.: 9,26e-304 Length: 626
Score: 3278.00 Matches: 625
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 91.4% Indels: 0
DB: 6 Gaps: 0

US-10-728-323-1 (1-2032) x ABU52572 (1-626)

QY 50 ATGAGAGGAGGGTTCTCCACTGATGCTGTGTAGGATCCTTGTCTCGCTTCAGTT 109
DB |||||||
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyIleLeuValLeuAlaSerVal 20
QY 110 TCTCAACGCGATGCCAAGTCATCACCTTACCAAGAAAGAAACAGAGAACCCCTCGCGCCAG 169
DB |||||||
DB 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGGTGCTCCAGAGTTGTCAACAGAACCGGATGACTTGAAGCAAAAGCATGCCAGTCT 229
DB |||||||
DB 41 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGCTGCACCAAGCTCGAGTATGATCCTCGTTGTGTCTATGATCCTCGAGGACACACTGGC 289
DB |||||||
DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCCAACCAACACGTTCCCTCCAGGAGCGGACACGTCGGCGCGCAACCCGAGACTAC 349
DB |||||||
DB 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATACCGCGCTCAACCCCGAAGAGAGAGGAGCGCGATGGGACCACTGGACCG 409
DB |||||||

Db 101 AspAspAlaArgArgGlnProArgArgGluGluGlyArgTrpGlyProAlaGlyPro 120
QY 410 AGGACCGTGAAGAGAAGAACTGGAGACAAACCAAGAGAAGATTGGAGGCGCAACACT 469
Db 121 ArgGluArgGluArgGluGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATCACAGCCACGGAAATAAGCCCGAAGGAGAGAGAGAACAAAGTGGGGAACA 529
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
QY 530 CCAGGTAGCCATGTGAGGGAAGAAACAATCTCGGAACAACCCCTTCTACTCTCCGTCGAAG 589
Db 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCCGCTACCGGAAACCAAAACGGTAGGATCCGGTCTCTGAGAGTTTGAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CARAGGTCAGGCAGGTTTCAGAAATCTCCAGAAATCCGTTATTTGTCAGATCCGAGGCCAAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAAACACTCTGTTCTTCCCAAGCACGCTGATGCTGATAACAATCTCTGTTATCCAGCAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGCACAGCCACCGTAGCCGAGCAAAATGGCAATTAACAGAAAGACTTTAATCTTGACGAG 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCATGCACTCAGAAATCCCATCCGTTTCATTTCTCTACATCTTGAAACCGCCATGACAAC 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACTCAGAGTAGTAAATCTCATGCCCGTTAAACACACCCGCCAGTTTGAGGAT 949
Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCAGCAGCCGAGACCAATCATCTACTTTCAGGGCTTCAGCGAATAACG 1009
Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGAGGCCCGCTTCAATCGGAATTCATGAGATACGAGGGTGTCTGTAGAGAGAAT 1069
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGTGAGCAAGAGGAGAGGCGAGCGATGAGTACTCGAGTAGTGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGAGTAGTAAAGTGTCAAAGGACGCTTGAAGAACTTACTAAGCAGCGCT 1189
Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCGCTCAAAGAGGCTCCGAGAGAGGAGATATCACCACCCCAACTCAACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGCGGAGCCCGATCTTCTTAACTTTGGGAAGTATTGTAGGTGAAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAACCCCGAGTTCCAGGACCTGGACATGATGCTCACTGTGTAGAGATCAAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCTGCTCGTCAACAAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAACCTGGAACCTTGAACTCGTGGCTGTAAAGAAAAGAGCAACAACAGAGGGGACGCGG 1489
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480

QY 1490 GAAGAAGGAGGACCAAGACCAAGAGAGGGAAGCTAAACAGAGAGGTGCGTAGGTAC 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTTCAAGGAAGGCGATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCCGAATCCATCTGCTGGCTTCGGTATCAACGCTGAAACCAACCAACACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTTCAGGTCATAAGACAACTGTGATAGACCATAGATAGACGAGGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTCGGGTCAACAAGTTGAGAAGCTCATCAAAAACCAAGGAATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTAGTGTCTCTCAATCTCAATCTCAATCTCAATCTCTCTCTCTCTCTCTCTCTCT 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TTCTCTCAGAAAGAGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
Db 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGGCTTTTAAC 1927
Db 621 IleLeuLysAlaPheAsn 626
RESULT 15
AAW22150
ID AAW22150 standard; protein; 626 AA.
XX AC AAW22150;
XX DT 17-OCT-2003 (revised)
DT 29-DEC-1997 (first entry)
XX Peanut allergen Ara hi.
DE Peanut; seed storage protein; allergen; allergy; hypersensitivity;
KW vaccine; anaphylactic shock; immunotherapy; therapy; monoclonal antibody;
KW ELISA; analysis; Ara hi.
XX Arachis hypogaea; strain Florunner.
XX Key Location/Qualifiers
FH Peptide 1..22
FT /label= Sig_peptide
FT Protein 23..626
FT /label= Mat_protein
FT Modified-site 521..523
FT /note= "N-glycosylation site"
XX WO9724139-A1.
XX 10-JUL-1997.
XX 23-SEP-1996; 96WO-US015222.
XX PF 29-DEC-1995; 95US-0009455P.
XX PR 04-MAR-1996; 96US-00610424.
XX (UYAR-) UNIV ARKANSAS.
XX Burks AW, Helm RM, Cockrell G, Stanley JS, Bannon GA;
PI WPI; 1997-363453/33.
XX N-PSDB; AAT76613.
DR

XX Peanut allergens Ara hi and Ara hii - used for vaccination and in two-
PT site monoclonal antibody based ELISA.
XX Claim 31, Page 172; 354pp; English.
XX This polypeptide comprises major peanut allergen Ara hi (AAW22149). Its
CC sequence was deduced from cDNA clone P41b (AA76613), isolated from
CC peanut seed cDNA using a primer (see AA76616) based on an isolated Ara
CC hi peptide (see AA24206). The sequence shows significant homology with
CC the vicilin family of seed storage proteins of other legumes. The
CC allergen is recognised by serum IgE from a large proportion of
CC individuals with peanut hypersensitivity. Ara hi and Ara hii (see
CC AA24164) can be used to raise monoclonal antibodies which are used in a
CC specific two-site MAB ELISA for the detection of Ara hi or Ara hii
CC (claimed). IgE-binding Ara hi antigen epitopes (see AA24165-87) may be
CC used in vaccines to protect against allergic reactions to peanut
CC allergens, e.g. anaphylactic shock. (Updated on 17-OCT-2003 to
CC standardise OS field)
XX SQ Sequence 626 AA;

Alignment Scores:
Pred. No.: 2,236-303 Length: 626
Score: 3274.00 Matches: 624
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 2
Query Match: 91.3% Indels: 0
DB: 2 Gaps: 0

US-10-728-323-1 (1-2032) x AAW22150 (1-626)

QY 50 ATGAGAGGAGGTTTCCACTGATGCTGTGTAGGATCCTGTCTCTGGCTTCAGTT 109
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyLeuValLeuAlaSerVal 20

QY 110 TTGCAACGCATGCCAAGTCATCCTTACCAGAAACACAGAACCCCTCGCCAC 169
DB 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40

QY 170 AGTGCTCCAGAGTTGTCAACAGAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCT 229
DB 41 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60

QY 230 CGTGTGACCAAGCTCCAGATATGATCCTCGTGTGTCTATGATCTCGAGGACACACTGC 289
DB 61 ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly 80

QY 290 ACCACCAACCAAGCTTCCCTCCAGGGAGCGGACACGTCGGCCGCAACCCGAGACTAC 349
DB 81 ThrThrAsnGlnArgSerProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100

QY 350 GATGATACCGCGCTCAACCCGAGAGAGGAGGCGCGATGGGACACGCTGGACCG 409
DB 101 AspAspAspArgArgGlnProArgGluGluGlyArgTyrGlyProAlaGlyPro 120

QY 410 AGGAGCGTGAAGAGCAAGAACTCGAGACCAACCAAGAGAGATTGGAGGCGACCAACT 469
DB 121 ArgGluArgGluGluGluAspTyrPargGlnProArgGluAspTyrPargProSer 140

QY 470 CATCAGCACCGGCAAAATAAGGCCGGAAGAGAGAGAGAGAAACAAGAGTGGGAAACA 529
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGluGluGlnGlnTyrGlyThr 160

QY 530 CCAGGTAGCCATGTGAGGAAAGAAACATCTCGAACAACCCCTTCTACTTCCCGTCAAGG 589
DB 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180

QY 590 CGGTTTAGCCCGCTACCGGAACCAAAACGGTAGGATCCGGTCTCTGAGAGGTTTGAC 649
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200

QY 650 CAAAGGTCAAGGCAGTTTCCAGAAATCTCCAGAAATCACCGGTATTGTGCAGATCGAGGCCAAA 709

DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTGTCTTCCCAAGCACGCTGATGCTGATAACATCTCTCTTATCCAGCAA 769
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTCACCGTACCAATGGCAATACAGAAAGAGCTTTAATCTTGACGAG 829
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GCCATGCACTCAGAAATCCATCCCGTTTCATCTTCTACATCTTGAACCGCATGACAAC 889
DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCCTCAGAGTAGCTAAATCTCCATGCCCGTTAAACACACCCGCGCAGTTGAGGAT 949
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValThrGlnProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGCAGCAGCGAGACCAATCATCTTCTGACGGGCTTCAGCAGGAATACG 1009
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGAGAGCGCCCTTCAATGCGGAATTCATAGATACGAGGGTGTCTGTAGAAGAGAT 1069
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluAsn 340
QY 1070 GCAGGAGGTGACCAAGAGAGAGGCGAGGCGGATCGGACTCGGAGTAGTGAGAAC 1129
DB 341 AlaGlyGlyGluGlnGluArgGlyGlnArgTyrPsrThrArgSerSerGluAsn 360
QY 1130 AATGAAGAGTGATAGTCAAAAGTGTCAAAGGAGCACGTTGAAGAAGCTTACTAAGCACGT 1189
DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCGTCTCAAAGAAAGGCTCCGAAGAGGAGATATACCAACCCATCAACTTG 1249
DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAGAGGCGGAGCCGATCTTCTAACCACTTTGGGAAGTATTATGAGGTGAGGCAGAC 1309
DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAGAAACCCCGACCTTCAGACCTGACATGATGCTCACCTGTGTAGAGATCAAGAA 1369
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTTGATGCTCCACACCTTCAACTCAAAAGGCGCATGGTTATCGTCTGCTCAACAA 1429
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
QY 1430 GGAAGTGGAAACCTTGAAGTCTGCTGCTNAGAAAAGACACACAGAGGGGCGCGG 1489
DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnArgGlyArgArg 480
QY 1490 GAAGAAGAGGAGGACCAAGACGAAGAGGAGGAGGAGTAACAGAGAGGTGCGTAGGTAC 1549
DB 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGGTGAAGAAAGGCGATGTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
DB 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AACGCTTCTCGAACTCCATCTGCTGGCTTCGGTATCAACGCTGAAAACACACACAGA 1669
DB 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTCTCTTGCAAGGTGATAAGGACAAATGTGATAGACCATAGACAGAACGAGGAT 1729
DB 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGTGCGGTGAACAAGTGTGAAGAGCTCATCAAAAACACAGAGGAATCT 1789

Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTCCTCGTCCTCAATCTCAATCTCCGTCGTCCTCGAGAAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerProSerProGluLysGlu 600
QY 1850 TCTCCTGAGAAAGAGGATCAAGAGGAGGAAAAACCAAGGAGGGTCCACTCCTTTCA 1909
Db 601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTGAAGCCTTTTAAC 1927
Db 621 IleLeuLysAlaPheAsn 626

Search completed: May 15, 2006, 21:45:53
Job time : 352.348 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: May 15, 2006, 22:18:32 ; Search time 12.2215 Seconds
(without alignments)
4123.811 Million cell updates/sec

Title: US-10-728-323-1
Perfect score: 3586
Sequence: 1 aataatcatatattcatc.....cgttgtgctgtttcttc 2032

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 572060 seqs, 82675679 residues
Total number of hits satisfying chosen parameters: 1144120
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p -model -DEV=xlp
-Q-/abs/ABSWEB_spool/US10728323/runat_15052006_172137_22463/app_query.fasta_1
-DB-issued_Patents_AA -QFMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs02p -USER=US10728323 @CGN 1.1.131 @runat_15052006_172137_22463
-NCPUS=6 -ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DRV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/iaa/5 COMB.pcp.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pcp.*
3: /cgn2_6/ptodata/1/iaa/H COMB.pcp.*
4: /cgn2_6/ptodata/1/iaa/PTUS COMB.pcp.*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pcp.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3278	91.4	626	2	US-09-106-872A-4 Sequence 4, Appli
2	1700	47.4	335	2	US-09-106-872A-17 Sequence 17, Appl
3	1279	35.7	605	1	US-07-955-905A-24 Sequence 24, Appl
4	1267	35.3	571	1	US-07-955-905A-25 Sequence 25, Appl
5	1171	32.7	449	2	US-10-245-227D-91 Sequence 91, Appl
6	1169	32.6	449	2	US-10-245-227D-95 Sequence 95, Appl
7	1169	32.6	449	2	US-10-245-227D-97 Sequence 97, Appl
8	1163	32.4	439	2	US-10-245-227D-1 Sequence 1, Appli
9	1163	32.4	449	2	US-10-245-227D-89 Sequence 89, Appl
10	1162	32.4	449	2	US-10-245-227D-93 Sequence 93, Appl
11	1157	32.3	417	2	US-10-245-227D-14 Sequence 14, Appl
12	1157	32.3	425	2	US-10-245-227D-12 Sequence 12, Appl

13	1117.5	31.2	390	2	US-10-245-227D-2	Sequence 2, Appli
14	1068	29.8	410	1	US-07-955-905A-26	Sequence 26, Appl
15	811	22.6	566	1	US-07-955-905A-2	Sequence 2, Appli
16	811	22.6	566	1	US-07-955-905A-22	Sequence 22, Appl
17	788	22.0	421	1	US-07-955-905A-27	Sequence 27, Appl
18	777	21.7	587	1	US-07-955-905A-23	Sequence 23, Appl
19	596.5	16.6	536	2	US-10-264-303-4	Sequence 4, Appli
20	595.5	16.6	540	2	US-10-264-303-3	Sequence 3, Appli
21	591.5	16.5	489	2	US-09-424-283-3	Sequence 3, Appli
22	569	15.9	524	2	US-09-424-283-1	Sequence 1, Appli
23	559.5	15.6	448	2	US-09-323-195A-18	Sequence 18, Appl
24	514.5	14.3	409	2	US-09-424-283-4	Sequence 4, Appli
25	513.5	14.3	523	2	US-09-323-195A-17	Sequence 17, Appl
26	501	14.0	444	2	US-09-424-283-2	Sequence 2, Appli
27	242.5	6.8	454	2	US-09-805-694B-4	Sequence 4, Appli
28	183	5.1	762	2	US-09-252-991A-29423	Sequence 29423, A
29	173	4.8	1564	2	US-10-144-198-2	Sequence 2, Appli
30	173	4.8	1564	2	US-10-144-198-4	Sequence 4, Appli
31	172	4.8	798	2	US-10-104-047-3335	Sequence 3335, Ap
32	166.5	4.6	720	2	US-09-252-991A-21881	Sequence 21881, A
33	164.5	4.6	1266	2	US-09-252-991A-10851	Sequence 10851, A
34	163	4.5	714	2	US-09-949-016-10296	Sequence 10296, A
35	160	4.5	630	2	US-09-252-991A-26324	Sequence 26324, A
36	160	4.5	801	2	US-09-252-991A-29274	Sequence 29274, A
37	160	4.5	1898	1	US-08-056-200-94	Sequence 94, Appl
38	160	4.5	1898	1	US-08-800-644-94	Sequence 1280, Ap
39	160	4.5	1898	2	US-09-538-092-1280	Sequence 1280, Ap
40	159	4.4	1043	2	US-09-252-991A-28885	Sequence 28885, A
41	158.5	4.4	1075	2	US-09-252-991A-18387	Sequence 18387, A
42	157	4.4	1780	1	US-08-769-309A-5	Sequence 5, Appli
43	157	4.4	1780	2	US-08-994-570-5	Sequence 5, Appli
44	157	4.4	1781	2	US-09-961-403-13	Sequence 13, Appl
45	157	4.4	2079	2	US-09-949-016-8301	Sequence 8301, Ap

ALIGNMENTS

RESULT 1
US-09-106-872A-4
Sequence 4, Application US/09106872A
Patent No. 6486311
GENERAL INFORMATION:
APPLICANT: Burks Jr., A. Weasley
APPLICANT: Stanley, J. Steven
APPLICANT: Cockrell, Gael
APPLICANT: King, Nina E.
APPLICANT: Sampson, Hugh A.
APPLICANT: Helm, Ricki M.
APPLICANT: Bannan, Gary A.
TITLE OF INVENTION: Peanut Allergens and Methods
FILE REFERENCE: HS 103 CIP
CURRENT APPLICATION NUMBER: US/09/106, 872A
CURRENT FILING DATE: 1999-06-29
PRIOR APPLICATION NUMBER: PCT/US96/15222
PRIOR FILING DATE: 1996-09-23
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 626
TYPE: PRT
ORGANISM: Arachis hypogaea
FEATURE:
OTHER INFORMATION: Amino Acids 25-34 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 1
OTHER INFORMATION: Amino Acids 48-57 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 2
OTHER INFORMATION: Amino Acids 65-74 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 3
OTHER INFORMATION: Amino Acids 89-98 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 4
OTHER INFORMATION: Amino Acids 97-106 are Ara H 1 binding epitope,
OTHER INFORMATION: peptide 5
OTHER INFORMATION: Amino Acids 107-116 are Ara H 1 binding epitope,

```

/ OTHER INFORMATION: peptide 6
/ OTHER INFORMATION: Amino Acids 123-132 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 7
/ OTHER INFORMATION: Amino Acids 134-143 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 8
/ OTHER INFORMATION: Amino Acids 143-152 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 9
/ OTHER INFORMATION: Amino Acids 294-303 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 10
/ OTHER INFORMATION: Amino Acids 311-320 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 11
/ OTHER INFORMATION: Amino Acids 325-334 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 12
/ OTHER INFORMATION: Amino Acids 344-353 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 13
/ OTHER INFORMATION: Amino Acids 393-402 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 14
/ OTHER INFORMATION: Amino Acids 409-418 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 15
/ OTHER INFORMATION: Amino Acids 461-470 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 16
/ OTHER INFORMATION: Amino Acids 498-507 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 17
/ OTHER INFORMATION: Amino Acids 525-534 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 18
/ OTHER INFORMATION: Amino acids 539-548 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 19
/ OTHER INFORMATION: Amino acids 551-560 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 20
/ OTHER INFORMATION: Amino acids 559-568 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 21
/ OTHER INFORMATION: Amino acids 578-587 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 22
/ OTHER INFORMATION: Amino acids 597-606 are Ara H 1 binding epitope,
/ OTHER INFORMATION: peptide 23
/ US-09-106-872A-4

Alignment Scores:
Pred. No.: 1,116-311 Length: 626
Score: 3278.00 Matches: 625
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 91.4% Indels: 0
DB: 2 Gaps: 0

US-10-728-323-1 (1-2032) x US-09-106-872A-4 (1-626)

QY 50 ATGAGAGGAGGTTTCTCCACTGATGCTGCTAGGATCCTGTCTCGGCTTCAGTT 109
DB 1 MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyLeuValLeuAlaSerVal 20

QY 110 TCTGCAACGATGCCAAGTCATCACCTTACCAGAAGAAACAGAGAACCCCTCGCCACG 169
DB 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40

QY 170 AGTGCCCTCAGAGTTGTCAACAGGAACCGATGACTTGAAGCAAAAGGCATCCGAGTCT 229
DB 41 ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer 60

QY 230 CGCTGGACCAAGCTCCAGATGATGCTCGTTGTGTCTATGATCCTCGAGGACACACTGGC 289
DB 61 ArgCysThrLysLeuGluLysArgPProArgCysValTyrAspProArgGlyHisThrGly 80

QY 290 ACCACCAACCAACGTTCCCTCCAGGGGACGACACAGTCGTCCGCCCAACCCGGAGACTAC 349
DB 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100

QY 350 GATGATGACCGCGTCAACCCGAGAGAGAGAGCGCGATGGGACCCAGCTGGACCG 409
DB 101 AspAspArgArgGlnProArgArgGluGluGlyArgTrpGlyProAlaGlyPro 120

QY 410 AGGAGCGTCAAAGAGAAGAACTGGAGACAACCAAGAGAAGATTGGAGCGCACCAAGT 469

```

```

DB 121 ArgGluArgGluArgGluArgGlnProArgTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATGACGAGCCACGAAATAGCCCGAAGAGAGAGAGAGAGACAGAGTGGGAAACA 529
DB 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnLutrpGlyThr 160
QY 530 CCAGTGTAGCATGTGAGGAGAAACATCTCGGAACAACCCCTTCTACTTCCGTCAGG 589
DB 161 ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCCGCTACGGGAACCAAAACCGTAGGATCCGGGTCTCTGAGAGGTTTAC 649
DB 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGGTCAAGCAGTTTCAGAAATCTCCAGAAATCAACCGTATTGTGCAGATCCAGGCCAAA 709
DB 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTCTTGTCTTCCCAAGCACGCTGATGCTGATAACATCTCTTATCCAGCAA 769
DB 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGCTACCGTAGCAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAG 829
DB 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCCATGCACTCAGAAATCCCATCCCGTTTCATTCTCTACATCTTGAACCCCATGCAAC 889
DB 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACCTCAGAGTAGTAAATCTCCATGCCGCTTAAACACACCCGCCAGTTTGAAGAT 949
DB 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCCGCGCAGCAGCCGAGACCAATCATCTACTTGCAGGGCTTCACAGGAATACG 1009
DB 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGluPheSerArgAsnThr 320
QY 1010 TTGGAGGCGCCTTCATCGGAAATCAATGATGATACGAGGGTGTCTTAGAGAGAT 1069
DB 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGGTGACCAAGAGAGAGAGCGCAGATGAGTACTCGGAGTAGTAGAGAAC 1129
DB 341 AlaGlyGlyGluGlnGluArgGlyGlnArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAGGAGTAGTAGTCAAAGTGTCAAAGGACACGTTGAAAGAACTTACTAAGCAGCT 1189
DB 361 AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla 380
QY 1190 AATCCGCTCAAAGAAAGCCTCCGAAAGAGAGGGAGATATCACCAACCCCACTCACTTG 1249
DB 381 LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 AGAAGAGGCGAGCCCGCATCTTTCTAAACAACCTTGGGAAGTATTTCAGGTGAAGCCAGAC 1309
DB 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAAACCCCGCTTCAGGACCTGGACATGATCTCACCTGTGTAGAGATCAAGAA 1369
DB 421 LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu 440
QY 1370 GAGCTTTGATGCTCCACACTTCAACTCAAGGCCATGGTTATCTCTCGTCGTCACAAA 1429
DB 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValValAsnLys 460
QY 1430 GGAACCTGGAAACCTTCAAACCTCGTGGCTGTAAAGAAAAAGAGCAACACAGAGGGGACGCGG 1489
DB 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnArgGlyArgArg 480
QY 1490 GAAGAGAGAGGACGCAAGACGAAGAGAGGAGGAAAGTAAACAGAGAGGTGCGTAGGTAC 1549
DB 481 GluGluGluGluAspGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 500

```

```
QY 1550 ACAGCGAGGTTGAAGGAGCGCATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
Db |||||||
501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AAGCTTCTCCGAACTCCATCTGCTGGTTCGGTATCAACGCTGCAAAAACACACAGAGA 1669
Db |||||||
521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCCTTGCAGTGCATGAAGACAATGTGATAGACCAGATAGAGAAAGCAAGCAGGAT 1729
Db |||||||
541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCATTCCTCGGGTCGGGTGAACAAAGTTGAGAACTCATCAAAAACAGAGGAATCT 1789
Db |||||||
561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCGTCTCGTCTCGTCAAGAGAG 1849
Db |||||||
581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerProGluLysGlu 600
QY 1850 TCTCCTGAGAAAGAGGATCAAGAGGAGGAGAAACCAAGGAGGAGGATCCACTCTTTTCA 1909
Db |||||||
601 SerProGluLysGluAspGlnGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGCTTTTAAC 1927
Db |||||||
621 IleLeuLysAlaPheAsn 626
```

RESULT 2

```
US-09-106-872A-17
; Sequence 17, Application US/09106872A
; Patent No. 6486311
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina E.
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricki M.
; APPLICANT: Bannion, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106,872A
; CURRENT FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: PCT/US96/15222
; PRIOR FILING DATE: 1996-09-23
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-09-106-872A-17
```

```
Alignment Scores:
Pred. No.: 2,28e-157 Length: 335
Score: 1700.00 Matches: 335
Percent Similarity: 99.4% Conservative: 0
Best Local Similarity: 99.4% Mismatches: 0
Query Match: 47.4% Indels: 2
DB: 2 Gaps: 1
```

```
US-10-728-323-1 (1-2032) x US-09-106-872A-17 (1-335)
```

```
QY 917 ATGCCCGTTAAACACACCCGGCCAGTTTGGAGTATCTCCCGGCGAGGAGCCGAGACCAA 976
Db |||||||
1 MetProValAsnThrProGlyGlnPheGluAspPhePheProAlaSerSerArgAspGln 20
QY 977 TCATCTACTTCAGGCGCTTCAGCAGGAATAGCTTCGAGGCGCGCTTCAATCGGAATTC 1036
Db |||||||
21 SerSerTyLeuGlnGlyPheSerArgAsnThrLeuGluAlaAlaPheAsnAlaGluPhe 40
```

```
QY 1037 AATGAGATACGAGGGTCTGTTAGAGAGAAATGCAGAGGTGACAAAGAGGAGAGGG 1096
Db |||||||
41 AsnGluIleArgArgValLeuLeuGluAsnAlaGlyGlyGluGlnGluArgGly 60
QY 1097 CAGAGCCGATGAGTACTCGGAGTACTGAGAACAAATGAAGGAGTATCTAACTGATCA 1156
Db |||||||
61 GlnArgArgTrpSerThrArgSerSerGluAsnAsnGluGlyValIleValLysValSer 80
QY 1157 AAGGAGACGCTTGAAGAAGCTTACTAAGCACGCTAAATCGTCTCAAAAGAAAGCTCCGAA 1216
Db |||||||
81 LysGluHisValGluGluLeuThrLysHisAlaLysSerValSerLysLysGlySerGlu 100
QY 1217 GAAGAGGAGATATACCAACCAATCACTTGAAGAGAGGAGCCGATCTTTCTTAAC 1276
Db |||||||
101 GluGluGlyAspIleThrAsnProIleAsnLeuArgGluGlyGluProAspLeuSerAsn 120
QY 1277 AACTTTGGGAGTATTTGAGGTCGAGCCAGCAAGAGAGAGCCCGAGCTTCAGGACCTG 1336
Db |||||||
121 AsnPheGlyLysLeuPheGluValLysProAspLysLysAsnProGlnLeuGlnAspLeu 140
QY 1337 GACATGATGCTCACCTGTGTAGAGATCAAAAGAGAGCTTTTGTGCTCCACACTTCAAC 1396
Db |||||||
141 AspMetMetLeuThrCysValGluLysGlyAlaLeuMetLeuProHisPheAsn 160
QY 1397 TCAAGGCCATGTTATCGTCTCTCAACAAAGAACTTGAACCTTGAACTCGTGGCT 1456
Db |||||||
161 SerLysAlaMetValIleValValAsnLysGlyThrGlyAsnLeuGluLeuValAla 180
QY 1457 GTAAGAAAGAGCAACAGAGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1516
Db |||||||
181 ValArgLysGluGlnGlnArgGlyArgArgGluGluGluGluAspGluAspGluGlu 200
QY 1517 GAGGAGGGAAGTAACAGAGAGGTGCTAGGTACACAGCGAGCTTCAAGAGAGCGCATGTG 1576
Db |||||||
201 GluGluGlySerAsnArgGluValArgArgTyThrAlaArgLeuLysGluGlyAspVal 220
QY 1577 TTCATCATGCCAGAGCTCATCCAGTCCATCAACGCTTCTCCGAACTCCATCTGCTT 1636
Db |||||||
221 PheIleMetProAlaAlaHisProValAlaIleAsnAlaSerSerGluLeuHisLeuLeu 240
QY 1637 GCTTTGGTATCAACGCTGAAACCAACACAGAAATCTTCTTGCAGGTGATAGGACAAT 1696
Db |||||||
241 GlyPheGlyIleAsnAlaGluAsnAsnHisArgIlePheLeuAlaGlyAspLysAspAsn 260
QY 1697 GTGATAGACCATGAGAGCAAGCAAGGAGTATTTAGCATTCCTTGGGTGGGTGAAACAA 1756
Db |||||||
261 ValIleAspGlnIleGluLysGlnAlaLysAspLeuAlaPheProGlySerGlyGluGln 280
QY 1757 GTTGAGAGCTCATCAAAAACAGAGGAATCTCACTTTGTGAGTCTCGTCTCAATCT 1816
Db |||||||
281 ValGluLysLeuIleLysAsnGlnLysGluSerHisPheValSerAla-----GlnSer 298
QY 1817 CAATCTCAATCTCGTCTCTCCAGAGAGAGTCTCTGAGAGAGAGTCTCTGAGAGAGAGATCAAGAGGAG 1876
Db |||||||
299 GlnSerGlnSerProSerSerProGluLysGluSerProGluLysGluAspGlnGluGlu 318
QY 1877 GAAACCAAGGAGGAGGCTCCACTCTCTTCAATTTTGAAGGCTTTTAAC 1927
Db |||||||
319 GluAsnGlnGlyGlyLysGlyProLeuLeuSerIleLeuLysAlaPheAsn 335
```

RESULT 3

```
US-07-955-905A-24
; Sequence 24, Application US/07955905A
; Patent No. 5770433
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 KD COCOA PROTEINS AND
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```

;
;   SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/955,905A
;   FILING DATE: 21-JAN-1993
;   CLASSIFICATION: 435
;   INFORMATION FOR SEQ ID NO: 24:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 605 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;     MOLECULE TYPE: protein
;     ORIGINAL SOURCE:
;     ORGANISM: Glycine max
;     FEATURE:
;     NAME/KEY: Protein
;     LOCATION: 1..605
;     OTHER INFORMATION: /note= "Vicilin from G. max"
;
; US-07-955-905A-24
;
Alignment Scores:
Pred. No.:      5,47e-116      Length:      605
Score:          1279.00        Matches:    285
Percent Similarity: 59.2%      Conservative: 107
Best Local Similarity: 43.1%    Mismatches: 174
Query Match:      35.7%        Indels:     96
DB:               1           Gaps:      18

US-10-728-323-1 (1-2032) x US-07-955-905A-24 (1-605)

QY 50 ATGAGAGGGAGGTTCTTCACCTGATGCTGCTAGGAGCTCTGCTGGCTTCAGTT 109
DB 2 MetArgAlaArgPhePro-----LeuLeuLeuGlyLeuValPheLeuAlaSerVal 19
QY 110 TCTGCAACGATGCCAAGTCATCACCTTACCAGAAGAAACACAGAACCCCTCGCCCG 169
DB 20 SerValSerPheGlyIleAla-----TyrTrpGluLys-----GluAsnProLysHisAsn 36
QY 170 AGTGCTCCAGAGTTGTCAACAGGAAACCGGATGACTTGAAGCAAAAGGATCGGAGTCT 229
DB 37 LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 56
QY 230 CGCTGCACCAAGCTCCAGTATGAT-----CCTCGTGT 262
DB 57 ArgCysAsnLeuLeuLysValGluLysGluLysGluLysGluLysGluLysGluLys 75
QY 263 GTCTATGATCTCGAGGACACACTGGCCACCACCAACCAACGTTCCCTCCAGGGAGCGG 322
DB 76 -----ProArgProArg-----ProGlnHisProGlu 84
QY 323 ACAGTGGCCGCCAACCCGAGACTACGATGATGAC----- 358
DB 85 ArgGluProGlnGlnProGlyGluLysGluLysAspGluAspGluAspGluProArgProIle 104
QY 359 -----CGCGTCAACCCCAAGAGAG-----GAAGGAGGCCGA 391
DB 105 ProPheProArgProGlnProArgGlnGluLysGluLysGluLysGluLysGluLys 124
QY 392 TGG-----GGACCACTGGACCGAGGAGCGCTGAAGAGAGAGAA 430
DB 125 TrpProArgLysGluGluLysArgGlyLysGlySerGluGluGluAspGluAspGlu 144
QY 431 GACTGGAGAACCAACAGAGAGAT-----TGGAGGCGACCAAGTCAATCAGCAGCCACGG 484
DB 145 AspGluGluGlnAspGluArgGlnPheProArgProProHisGlnLysGluGlu 164
QY 485 AAATAAGGCCCGAAGGAGAGAGAGAACAGAGTGGGAAACACCAAGTACGATGTC 544
DB 165 ArgAsnGluGluAspGluAspGluGlnGlnGlnArgGluSerGluSerGluAsp 184
QY 545 AGGGA-----GAAACATCTCGGAACACCTTTCTACTTCCCGTCAAGCGCGTTTACG 598
DB 185 SerGluLeuArgArgHisLysAsnLysAsnProPheLeuPhePheSerAsnArgPheGlu 204
```

```

QY 599 ACCCGCTACGGGAACCAAAACGGTAGGATCCGGGTCTCTGACAGAGTTTTCACCAAGGTCA 658
DB 205 ThrLeuPheLysAsnGlnTyrGlyArgIleArgValLeuGlnArgPheAsnGlnArgSer 224
QY 659 AGCGAGTTTTCAGAATCTCCAGAAATCACCGTATTTGTGCAGATCGAGGCCAAACCTTAACACT 718
DB 225 ProGlnLeuGlnAsnLeuArgAspTyrArgIleLeuGluPheAsnSerLysProAsnThr 244
QY 719 CTGTGTTCTTCCAAGCACGCTGATGCTGATAACATCTCTTGTATTATCCAGCAAGGGCAAGCC 778
DB 245 LeuLeuLeuProAsnHisAlaAspAlaAspTyrLeuIleValIleLeuAsnGlyThrAla 264
QY 779 ACCGTACCGTAGCAAAATGCGCAATACAGAAAGAGCTTTAATCTTCACGAGGCCCATGCA 838
DB 265 IleLeuSerLeuValAsnAsnAspArgAspSerTyrArgLeuGlnSerGlyAspAla 284
QY 839 CTCAGAAATCCCATCCGGTTTTCATTTCTTACATCTTTGAACCGCATGACACCAACGACCTC 898
DB 285 LeuArgValProSerGlyThrThrTyrTyrValValAsnProAspAsnAsnGluAsnLeu 304
QY 899 AGAGTAGCTAAATCTCCATGCCCCTTAACACACCCGGCCAGTTTGTAGGATTTCTTCCCG 958
DB 305 ArgLeuIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSerPhePheLeu 324
QY 959 GCGAGCGCGAGACCAATCATCTACTTTCAGGGCTTCAGCAGGAATACGTTGGAGGCC 1018
DB 325 SerSerThrGluAlaGlnGlnSerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAla 344
QY 1019 GCCTTCAATCGCGAATTCATAGATACGAGAGGGTCTCTTTAGAAGAGATCGCAGAGGT 1078
DB 345 SerTyrAspThrLysPheGluGluIleAsnLysValLeuPheSerArgGluGluGly--- 363
QY 1079 GAGCAAGAGAGAGAGAGCGGAGCGATCGAGTACTCGGAGTAGTGAGACAATGAGGA 1138
DB 364 -----GlnGlnGlnGlyGlnArgLeu-----GlnGluSer 374
QY 1139 GTGATAGTCAAGTGCAAGGAGCAGCTTTGAAGAACTTACTTAAGCACGCTAAATCCGTC 1198
DB 375 ValIleValGluLysSerLysGluGlnIleArgAlaLeuSerLysArgAlaLysSerSer 394
QY 1199 TCAAGAAAGGCTCCCAAGAGAGGAGATATCACCAACCCCAATCAACTTGAGAGAGGC 1258
DB 395 SerArgLysThrIleSerSerGlu-----AspLysProPheAsnLeuArgSerArg 411
QY 1259 GAGCCGCTTCTTCTTAACAACTTTGGGAAGTTATTGAGTGAGGAGCCACACAGAGAAC 1318
DB 412 AspProIleTyrSerAsnLysLeuGlyLysPhePheGluIleThrProGlu---LysAsn 430
QY 1319 CCCAGCTTCAGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAAGAGAGCTTTG 1378
DB 431 ProGlnLeuArgAspLeuAspIlePheLeuSerIleValAspMetAsnGluGlyAlaLeu 450
QY 1379 ATGCTCCCACTTCAACTCAAAGGCCATGGTTATCGTCTCGTCAACAAAGAACTGGA 1438
DB 451 LeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGlyAspAla 470
QY 1439 AACCTTGAATCTGCTGTGAAGAAAGACAACAGAGGGGAGCGGGGAGAGAGAG 1498
DB 471 AsnIleGluLeuValGlyLeu---LysGluGlnGlnGlnGlnGlnGlnGlnGln 499
QY 1499 GAGGACGAGACGAAAGAGGAGGAAAGTAAACAGAGAGGTGCGTAGGTACACAGCGAGG 1558
DB 490 Pro-----LeuGluValArgLysTyrArgAlaGlu 499
QY 1559 TTGAAGGAAGGCGATGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCC 1618
DB 500 LeuSerGluGlnAspIlePheValIleProAlaGlyTyrProValValValAsnAlaThr 519
QY 1619 TCCGAATCCATCTGCTGCTTGGCTTCAACGCTGAAACCAACCAACCAACCAATCTTCCCT 1678
DB 520 SerAsnLeuAsnPhePheAlaIleGlyIleAsnAlaGluAsnAsnGlnArgAsnPheLeu 539
QY 1679 GCAGGTGATAAGGACAATGTGTAGACCATAGAGAAGCAAGCAAGGAGTTTATGACATTC 1738
```

Db 540 AlaGlySerGlnAspAsnValIleSerGlnIleProSerGlnValGlnGluLeuAlaPhe 559
| | | | | : : : | | | | | | | | | | | | | | | : : : : : | | | | |
Qy 1739 CCTGGCTCGGTGAACAAGTTGAGAGCTCATCAAAAACAGAGGAATCTCACTTTGTG 1798
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 560 ProGlySerAlaGlnAlaValGluLysLeuLeuLysAsnGlnArgGlySerTyrPheVal 579
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 1799 AGTGCTCGCTCAATCTCAATCTCAATCTCCGCTCTCCTGAGAAAGAGTCTCCTGAG 1858
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 580 AspAlaGlnProLysLys 585
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 1859 AAGAGGATCAAGAGAGGAAACCAAGAGGAGGAGGTCCTCACTCTTTCAATTTTGAAG 1918
: : | | | | | | | | | | | | | | | | | | | | | | | | | : : : : :
Db 586 -----LysGluGluGluGlyAsnLysGlyArgLysGlyProLeuSerSerIleLeuArg 602
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 1919 GCTTTT 1924
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 603 AlaPhe 604

RESULT 4

US-07-955-905A-25
; Sequence 25 Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 kD COCOA PROTEINS AND
; TITLE OF INVENTION: PRECURSOR
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 571 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Pium sativum
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..571
; OTHER INFORMATION: /note= "Conviciilin from P. sativum"
US-07-955-905A-25

Alignment Scores:

Pred. No.: 7,88e-115 Length: 571
Score: 1267.00 Matches: 276
Percent Similarity: 56.7% Conservative: 98
Best Local Similarity: 41.8% Mismatches: 146
Query Match: 35.3% Indels: 140
DB: 1 Gaps: 13

US-10-728-323-1 (1-2032) x US-07-955-905A-25 (1-571)

Qy 68 CCACTGATGCTGTGCTAGGGATCCTTGTCTCGGCTTCAGTTCTGCAACGATGCGAAG 127
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 10 ProLeuLeuLeuPheLeuGlyIleIlePheLeuAlaSerValCysValThrTyrAlaAsn 29
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 128 TCATCACTTACAGAAAGAAACAGAGAACCCCTGCGCCAGAGGTGCTCCAGAGTTGT 187
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 29 ----- 29
Qy 188 CAACAGGAACCGATGACTTTGAAGCAAAAGGATGCGAGTCTCGTGCACCAAGCTCGAG 247
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 29 ----- 29

Qy 248 TATGATCCTCGTTGTCTATGATCCTCGAGGACACACTGGCACCACCACCAACCAACGTTCC 307
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 30 -----TyrAspGluGlySerGluThrArgVal----- 38
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 308 CCTCAGGGGACGACACAGTGGCGCCGACCAACCCGAGAGACTAGATGATGATCAGCCCGTCAA 367
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 39 ---ProGlyGlnArgGluArgGlnGluGlyGluGlyGluGluLysArgHis--- 56
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 368 CCCCAAGAGAGAGGAGGCGATGGGACCCAGCTGGACCCGAGGAGCGT----- 418
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 57 -----GlyGluTrpArgProSerTyrGluLysGluGluHisGluGlu 70
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 418 ----- 418
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 71 GluLysGlnLysTyrArgTyrGlnArgGluLysLysGluGlnLysGluValGlnProGly 90
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 419 -----GAAAGAGAGAGAGACTGGAGACACCAACCAAGAGAGATTGGAGG----- 460
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 91 ArgGluArgTrpGluArgGluAspGluGlnValGlnGluGluTrpArgGlySer 110
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 461 -----CGACCAAGTCATCAGCAGGCCACCG 484
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 111 GlnArgArgGluAspProGluGluArgAlaArgLeuArgHisArgGluGluArgThrLys 130
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 485 AAAATAAGGCCCCGAAAGAGAGAGAGAGAGAGAGTGGGGAACACAGGTAGCCATGTG 544
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 131 ArgAspArgArgHisGlnArgGluGlyGluGluGlu-----GluArgSer 145
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 545 AGGGAAGAAACATCTCGGAACCAACCTTTCTACTTCCCGTCAAGCGGTTTACACCCCG 604
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 146 SerGluSerGlnGluHisArgAsnProPheLeuPheLysSerAsnLysPheLeuThrLeu 165
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 605 TACGGGAACCAAAACGCTAGGATCCGGGCTCTCGAGAGGTTTGACCAAGGTTCAAGGCAG 664
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 166 PheGluAsnGluAsnGlyHisIleArgArgLeuGlnArgPheAspLysArgSerAspLeu 185
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 665 TTTGAGAATCTCCAGAATCACCGTATTGTGAGATCGAGGCCAAACCTCAACTCTTTGT 724
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 186 PheGluAsnLeuGlnAsnTyrArgLeuValGluTyrArgAlaLysProHisThrIlePhe 205
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 725 CTTCCCAAGCAGCTGATGCTGATTAACATCCTCTTATCCAGCAAGGCAAGCCCGTG 784
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 206 LeuProGlnHisIleAspAlaAspLeuIleLeuValLeuAsnGlyLysAlaIleLeu 225
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 785 ACCGTAGCAAAATGCAATAACAGAAAGAGCTTTAATCTTGACGAGGCCATGCACCTCAGA 844
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 226 ThrValLeuSerProAsnAspArgAsnSerTyrAsnLeuGluArgGlyAspThrIleLys 245
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 845 ATCCCATCCGGTTTCATTTCTTACATCTTGAAACCGCCATGACACACAGAACTCAGAGTA 904
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 246 IleProAlaGlyThrThrSerTyrLeuValAsnGlnAspAspGluGluAspLeuArgVal 265
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 905 GCTAAATCTCCATCCCGTTAAACACACCCCGCCAGTTTGAGATTCTTCCCGGCGAGC 964
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 266 ValAspPheValIleProValAsnArgProGlyLysPheGluAlaPhe-----GlyLeu 283
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 965 AGCCGAGACCAATCATCTTCTTACATCTTGAGAGGCTTTCAGCAGGAATACGTTGGAGCGCGCTTC 1024
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 284 SerGluAsnLysAsnGlnTyrLeuArgGlyPheSerLysAsnIleLeuGluAlaSerLeu 303
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 1025 AATCGGGAATTCATAGATACGAGGCTGTGTTAGAGAGAAATTCAGGAGGTGAGCAA 1084
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 304 AsnThrLysTyrGluThrIleGluLysValLeuLeuGluGlnGluLysLysProGln 323
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 1085 GAGGAGAGAGGCGACAGGCGATGAGTACTCGAGTAGTAGAGAACTAAGAGAGTGATA 1144
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 324 GlnLeuArgAspArgLysArg-----ThrGlnGlnGlyGluGluArgAsp---AlaIle 340
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 1145 GTCAAGTGTCAAAGGACGACGCTTCAAGAACTTAAAGCAGCTTAAATCCGCTCTCAAAG 1204
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 341 IleLysValSerArgGluGlnIleGluLeuArgLysLeuAlaLysSerSerLys 360
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Qy 1205 AAAGGCTCCGAAGAGAGGAGATATACCAACCCCAATCACTTGAGAGAGCGGAGCCC 1264
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |

```
Db      361  LysSerLeuProSerGluPheGlu-----ProPheAsnLeuArgSerHisLysPro 377
      1265  GATCTTTCTAAACAACTTTGGGAAGTTATTGAGGTGAAGCAGACACAAAGAAACCCCCAG 1324
      378  GluTyrSerAsnLysPheGlyLysLeuPheGluLeuThrProGluLysLysTyrProGln 397
      1325  CTTCAGAGCTGACATGATGCTCACCCTGTGTAGAGATCAAGAAGAGAGCTTTGTGCTC 1384
      398  LeuGlnAspLeuAspIleLeuValSerCysValGluIleAsnLysGlyAlaLeuMetLeu 417
      1385  CCACACTTCAACTCAAGGCCATGGTTATCGTCGTCTCAACAAAGAACTGGAACCTT 1444
      418  ProHisTyrAsnSerArgAlaIleValValLeuLeuValAsnGluGlyLysGlyAsnLeu 437
      1445  GAACTCGTCTGCTTAAGAAAAGAGCAACACAGAGGGGACGGCGGAAGAAGAGGAGAC 1504
      438  GluLeuLeuGlyLeuLysAsnGluGlnGluArg-----449
      1505  GAAGACGAAGAAGAGAGGAGGAGTAACACAGAGAGGTGGTAGGTACACAGGAGGTGAAG 1564
      450  GluAspArgLysGluArg-----AsnAsnGluValGlnArgTyrGluAlaArgLeuSer 467
      1565  GAAGCGCATGTCTCATCGCCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAA 1624
      468  ProGlyAspValValIleIleProAlaGlyHisProValAlaIleSerAlaSerAsn 487
      1625  CTCATCTGTGCTGGCTTCGCTATCAACGCTGAAACAAACACACAGAAATCTTCTTCG 1684
      488  LeuAsnLeuLeuGlyPheGlyIleAsnAlaLysAsnAsnGlnArgAsnPheLeuSerGly 507
      1685  GATAAGACAAATGTATAGACCGATAGAGAAAGCAAGCGAAGATTAGCATTTCCCTGG 1744
      508  SerAspAspAsnValIleSerGlnIleGluAsnProValLysGluLeuThrPheProGly 527
      1745  TCGGGTCAACAGTTGAGAGCTCATCAAAACCAAGAACTCACTTTGTGAGTGCT 1804
      528  SerSerGlnGluValAsnArgLeuIleLysAsnGlnLysGlnSerHisPheAlaSerAla 547
      1805  CGTCTCAATCTCAATCTCAATCTCCGCTCGTCTCTCTGAGAAAGAGTCTCTCGAGAAG 1864
      548  GluPro-----Glu 550
      1865  GATCAGAGAGGAGAAACCAAGAGGAGGAGGTCCTCTTCAATTTTGAAGGCTTTT 1924
      551  GlnLysGluGluGluSerGlnArgLysArgSerProLeuSerValLeuAspSerPhe 570
```

RESULT 5

```
US-10-245-227D-91
; Sequence 91, Application US/10245227D
; Patent No. 6936696
; GENERAL INFORMATION:
; APPLICANT: Monsanto Company
; APPLICANT: Rapp, William D.
; APPLICANT: Peng, JieXin
; APPLICANT: Nadiq, Gautham
; APPLICANT: Venkatesh, Tyamagondlu
; TITLE OF INVENTION: Enhanced Protein and Methods of Their Use
; FILE REFERENCE: REN-00-087 US
; CURRENT APPLICATION NUMBER: US/10/245,227D
; CURRENT FILING DATE: 2002-09-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 91
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Glycine max
US-10-245-227D-91
```

```
Alignment Scores:
Pred. No.: 1.69e-105 Length: 449
Score: 1171.00 Matches: 239
Percent Similarity: 70.4% Conservative: 87
```

```
Best Local Similarity: 51.6% Mismatches: 83
Query Match: 32.7% Indels: 54
DB: 2 Gaps: 10
US-10-728-323-1 (1-2032) x US-10-245-227D-91 (1-449)
QY      542  GTGAGGAGAAACATCTCGGAACACCCCTTTCTACTTC---CCGTCAAGCGGTTTAGC 598
Db      36  ValArgGluAsp-----GluAsnAsnProPheTyrLeuArgSerSerAsnSerPheGln 53
QY      599  ACCCGCTACGGGAACCAAAACCGTAGGATCCGGGCTCTCGAGAGGTTTGCACCAAGGTCA 658
Db      54  ThrLeuPheGluAsnGlnAsnGlyArgIleArgLeuIleGlnArgPheAsnLysArgSer 73
QY      659  AGCGAGTTTCAGAAATCCAGAAATCACCGTATTGTGCAGATCGAGGCCAACCTAACACT 718
Db      74  ProGlnLeuGluAsnLeuArgAspTyrArgIleValGlnPheGlnSerLysProAsnThr 93
QY      719  CTTGTCTTCTCCCAAGCACCGCTGATGCTGATAATCATCTTGTATTATCCAGCAAGGCGCAAGCC 778
Db      94  IleIleLeuProHisHisAlaAspAlaAspPheIleIleIleIleLeuSerGlyArgAla 113
QY      779  ACCGTACCGTAGCAAAATGGCAATACAGAAAAGAGCTTTAATCTTCACGAGGCGCCATGCA 838
Db      114  IleLeuThrLeuValAsnAsnAspAspArgAspSerTyrAsnIleHisProGlyAspAla 133
QY      839  CTCAGAAATCCCATCCCGTTTCATTTCTCATCTTGAACCGGCATCAACACCAAGACCTC 898
Db      134  GlnArgIleProAlaGlyThrThrTyrTyrIleIleAsnProHisAspHisGlnAsnLeu 153
QY      899  AGAGTAGCTAAATCTCCATGCCGTTTAAACACACCCGCGCAGTTTCAGGAGTTTCTTCCCG 958
Db      154  LysIleIleLysIleAlaIleProValAsnLysProGlyArgTyrAspPhePheLeu 173
QY      959  GCGAGCAGCGCAGACCAATCATCTTCTGCGGGCTTCAGCAGGAATACGTTGGAGGCC 1018
Db      174  SerSerThrGlnAlaGlnGlnSerTyrLeuGlnGlyPheSerHisAsnIleLeuGluThr 193
QY      1019  GCCTTCAATCGGAATTCATAGATACGAGAGTCTGTTAGAGAGATCGAGAGGT 1078
Db      194  SerPheHisSerGluPheGluGluIleAsnArgValLeuLeu-----Gly 208
QY      1079  GAGCAGAGAGAGGCGAGGCGATGGAGTACTCGGAGTAGTCAGAACAAATGAAGGA 1138
Db      209  GluGluGluGluGlnArgGlnGln-----GluGly 218
QY      1139  GTGATAGTCAAAGTGTCAAAGGAGCAGTTTGAAGAACTTACTAAGCAGCTAAATCCGTC 1198
Db      219  ValIleValGluLeuSerLysGluGlnIleArgGlnLeuSerArgArgAlaLysSerSer 238
QY      1199  TCAAAGAAA-----GGCTCCGAGAGAGGAGGAGATATCACCAACCAATCAACTTGAGA 1252
Db      239  SerArgLysThrIleSerSerGluAspGlu-----ProPheAsnLeuArg 253
QY      1253  GAAGGCGAGCCCGATCTTTCTAACTTTGGGAAAGTTATTTGAGGTGAAGCCAGCAAG 1312
Db      254  SerArgAsnProIleTyrSerAsnAsnPheGlyLysPheGluIleThrProGlu--- 272
QY      1313  AAGAAACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAGGA 1372
Db      273  LysAsnProGlnLeuArgAspLeuAspIlePheLeuSerSerValAspIleAsnGluGly 292
QY      1373  GCCTTGTATGCTCCACACITCAACTCAAAGGCCATGGTTATCTGCTCGTCAACAAAGGA 1432
Db      293  AlaLeuLeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGly 312
QY      1433  ACTGGAAACCTTGAACTCGTGGCTGTAAAGAAAAGAGCAACACAGAGGGGACCGCGGAA 1492
Db      313  AspAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnGlnLysGlnLysGlnGlu 331
QY      1493  GAAGAGAGGAGCAAGACGAAGAAGAGGAGGAGGAAGTAAACAGAGAGGTGCTAGGTACACA 1552
Db      332  GluGluPro-----LeuGluValGlnArgTyrArg 341
```


; APPLICANT: Peng, Jixian
; APPLICANT: Nadig, Gautham
; APPLICANT: Venkatesh, Tyamagondlu
; TITLE OF INVENTION: Enhanced Protein and Methods of Their Use
; FILE REFERENCE: REN-00-087 US
; CURRENT APPLICATION NUMBER: US/10/245,227D
; CURRENT FILING DATE: 2002-09-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Glycine max
US-10-245-227D-1

Alignment Scores:
Pred. No.: 1,01e-104 Length: 439
Score: 1163.00 Matches: 238
Percent Similarity: 70.2% Conservative: 87
Best Local Similarity: 51.4% Mismatches: 84
Query Match: 32.4% Indels: 54
DB: 2 Gaps: 10

US-10-728-323-1 (1-2032) x US-10-245-227D-1 (1-439)

```
QY 542 GTGAGGGAAGAAACATCTCGGAACACCCCTTTCTACTTC---CCGTCAAGGGCGTTTAGC 598
DB 26 ValArgGluAsp-----GluAsnAsnProPheTyrLeuArgSerSerAsnSerPheGln 43
QY 599 ACCCGTACGGGAACCAAAACGGTAGGATCCGGGCTCGAGAGTTTGACCAAGGTCA 658
DB 44 ThrLeuPheGluAsnGlnAsnGlyArgIleArgLeuGlnArgPheAsnLysArgSer 63
QY 659 AGGCAGTTTCAGAACTTCAGAAATCACCGTATTGTGCAGATCGAGGCCAACTACACT 718
DB 64 ProGlnLeuGluAsnLeuArgAspTyrArgIleValGlnPheGlnSerLysProAsnThr 83
QY 719 CTTGTTCTTCCCAAGCACGCTGATGCTGATAACATCTTGTATTATCCAGCAAGGGCAAGCC 778
DB 84 IleLeuLeuProHisHisIleAspAlaAspPheLeuLeuPheValLeuSerGlyArgAla 103
QY 779 ACCGTACCGTACCAATGCGCAATACAGAAAGAGCTTTAATCTTACGAGGCCATGCA 838
DB 104 IleLeuThrLeuValAsnAsnAspArgAspSerTyrAsnLeuHisProGlyAspAla 123
QY 839 CTCAGAAATCCATCCCGTTTCATTTCTCATCTTGAACCGCCATCACACCAAGACCTC 898
DB 124 GlnArgIleProAlaGlyThrThrTyrTyrLeuValAsnProHisAspHisGlnAsnLeu 143
QY 899 AGAGTAGTAAATCTCCATGCCCGTTAAACACACCGCGCCAGTTTGAGGATTTCTTCCCG 958
DB 144 LysIleLeuLeuAlaIleProValAsnLysProGlyArgTyrAspAspPhePheLeu 163
QY 959 GCGAGACGGGACCAATCATCTCTTACGGGCTTCAGAGGAATACGTTGGAGGCC 1018
DB 164 SerSerThrGlnAlaGlnGlnSerTyrLeuGlnGlyPheSerHisAsnIleLeuGluThr 183
QY 1019 GCCTCAATCGGAATTCATAGATACGAGGGTCTCTTGAAGAGATCGAGAGGT 1078
DB 184 SerPheHisSerGluPheGluGluIleAsnArgValLeuLeu-----Gly 198
QY 1079 GAGCAAGAGAGAGGCGAGGCGATCGAGTACTCGGAGTAGTGAGAACAAATGAAGA 1138
DB 199 GluGluGluGluGlnArgGln-----GluGly 208
QY 1139 GTGATAGTCAAGTGTCAAGAGACAGCTTTGAGAACTTACTAAGCACCGCTAAATCCGTC 1198
DB 209 ValIleValGluLeuSerLysGluGlnIleArgGlnLeuSerArgArgAlaLysSerSer 228
QY 1199 TCAAGAAA-----CGCTCCGAGAGAGGGGAGATATCACCAACCAATCACTTGAGA 1252
DB 229 SerArgLysThrIleSerSerGluAspGlu-----PropheAsnLeuArg 243
```

```
QY 1253 GAAGGCGAGCCCGATCTTTTAAACAATTTGGAAAGTTATTATTGAGTGAAGCCAGACAAG 1312
DB 244 SerArgAsnProIleTyrSerAsnAsnPheGlyLysPheGluIleThrProGlu--- 262
QY 1313 AAGAAACCCCGACCTTCAGACCTGACATGATGCTCACCTGTGTAGAGATCAAGAGGA 1372
DB 263 LysAsnProGlnLeuArgAspLeuAspIlePheLeuSerSerValAspIleAsnGluGly 282
QY 1373 GCTTTGATGCTCCCACTTCAACTCAAAGGGCCATGTTATCTGCTCCTCAACAAGGA 1432
DB 283 AlaLeuLeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGly 302
QY 1433 ACTGAAACCTTGAATCTGCTGTAAAGAAAGCAACAACAGAGGGGACGCGGGAA 1492
DB 303 AspAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnGlnLysGlnGlu 321
QY 1493 GAAGAGGAGGACGAAGCAAGAGGAGGAAGTAAACAGAGAGGTGCGTAGGTACACA 1552
DB 322 GluGluPro-----LeuGluValGlnArgTyrArg 331
QY 1553 GCGAGGTGAAGGAAGGCGATGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAAC 1612
DB 332 AlaGluLeuSerGluAspAspValPheValIleProAlaIleTyrProPheValValAsn 351
QY 1613 GCTTCTCCGAATCCATCTGCTTGGCTTCCGATATCAACGCTGAAACCAACACAGAATC 1672
DB 352 AlaThrSerAsnLeuAsnPheLeuAlaPheGlyIleAsnAlaGluAsnGlnArgAsn 371
QY 1673 TTCCTTGCAGGTGATAAGGACAAATGTATAGACCAAGATAGAGAACCAAGCAAGGATTTA 1732
DB 372 PheLeuAlaGlyGluLysAspAsnValValArgGlnIleGluArgGlnValGlnGluLeu 391
QY 1733 GCATTCCTCTGGTCCGGTGAACAAGTTGAGAAGCTCATCAAAACCAAGAGGAATCTCAC 1792
DB 392 AlaPheProGlySerAlaGlnAspValGluArgLeuLeuLysGlnArgGluSerTyr 411
QY 1793 TTTGTAGTGTCTGCTCCTCAATCTCAATCTCCTGCTCCTGCTGAGAAAGAGTCT 1852
DB 412 PheValAspAlaGlnProGln----- 418
QY 1853 CCGTGAAGAGAGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1912
DB 419 -----GlnLysGluGluGlySerLysGlyArgLysGlyProPheProSerIle 434
QY 1913 TTGAAGGCT 1921
DB 435 LeuGlyAla 437
```

RESULT 9

US-10-245-227D-89
; Sequence 89, Application US/10245227D
; Patent No. 6936696
; GENERAL INFORMATION:
; APPLICANT: Monsanto Company
; APPLICANT: Rapp, William D.
; APPLICANT: Peng, Jixian
; APPLICANT: Nadig, Gautham
; APPLICANT: Venkatesh, Tyamagondlu
; TITLE OF INVENTION: Enhanced Protein and Methods of Their Use
; FILE REFERENCE: REN-00-087 US
; CURRENT APPLICATION NUMBER: US/10/245,227D
; CURRENT FILING DATE: 2002-09-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 89
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Glycine max
US-10-245-227D-89

Alignment Scores: 1.03e-104 Length: 449
Pred. No.: 1163.00 Matches: 238
Score:

```
Percent Similarity: 70.2%      Conservative: 87
Best Local Similarity: 51.4%    Mismatches: 84
Query Match: 32.4%             Indels: 54
DB: 2                           Gaps: 10

US-10-728-323-1 (1-2032) x US-10-245-227D-89 (1-449)

QY 542 GTGAGGAGAAACATCTCGGAACAAACCTTTTACTTC---CCGTCAAGCGGTTTAGC 598
Db 36 ValArgGluAsp-----GluAsnAsnProPheTyrLeuArgSerSerAsnSerPheGln 53
QY 599 ACCCGCTAGCGGAACCAAAACGCTAGGATCCGGTCTCTGAGAGGTTTTCACCAAGGTCA 658
Db 54 ThrLeuPheGluAsnGlnAsnGlyArgGileArgLeuGlnArgPheAsnLysArgSer 73
QY 659 AGGCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAACCTAACACT 718
Db 74 ProGlnLeuGluAsnLeuArgAspTyrArgIleValGlnPheGlnSerLysProAsnThr 93
QY 719 CTTGTTCTCCCAAGCAGCGTGTATGCTGATAACATCTTGTGTATCCAGCAGGCGAAGCC 778
Db 94 IleLeuLeuProHisHisAlaAspPheLeuLeuPheValLeuSerGlyArgAla 113
QY 779 ACCTGACCGTACCAATGCGCAATAACAGAAAGAGCTTTAATCTTCAGCAGGCGCATGCA 838
Db 114 IleLeuThrLeuValAsnAsnAspArgAspSerTyrAsnLeuHisProGlyAspAla 133
QY 839 CTCAGAAATCCATCCCGTTTCATTTCTCATCTTGAACCGCATGACACCAAGAACCTC 898
Db 134 GlnArgIleProAlaGlyThrTyrTyrLeuValAsnProHisAspHisGlnAsnLeu 153
QY 899 AGAGTAGCTAAATCTCCATGCCCGTTTAAACACACCCGCGCAGTTTGAGGATTTCTCCCG 958
Db 154 LysIleLeuLysLeuAlaIleProValAsnLysProGlyArgTyrAspAspPhePheLeu 173
QY 959 GCGAGCAGCGGACCAATCATCTCTTCGCGGCTTCAGAGGAAATACGTTTGGAGGCC 1018
Db 174 SerSerThrGlnAlaGlnSerTyrLeuGlnGlyPheSerHisAsnIleLeuGluThr 193
QY 1019 GCCTTCAATCGGAATTCATAGATACGAGGAGTGTCTGTAGAAGAGATCGAGGAGGT 1078
Db 194 SerPheHisSerGluPheGluGluIleAsnArgValLeuLeu-----Gly 208
QY 1079 GACCAAGAGAGAGAGGCGAGCGAGTGGAGTACTCGGAGTAGTGAGAACATGAAGGA 1138
Db 209 GluGluGluGluGlnArgGlnGln-----GluGly 218
QY 1139 GTGATAGTCAAGTGTCAAAGGACGCTTGAAGAACTTACTAAGCAGCGCTAAATCGCT 1198
Db 219 ValIleValGluLeuSerLysGluGlnIleArgGlnLeuSerArgArgAlaLysSerSer 238
QY 1199 TCAAGAAA-----GCTCCGAGAGAGGAGAGATATCAACCAACCAATCAACTTGAGA 1252
Db 239 SerArgLysThrIleSerSerGluAspGlu-----ProPheAsnLeuArg 253
QY 1253 GAAGGCGAGCCCATCTTCTTCAACATTTGGGAAGTTATTTCAGGTGAGCCAGACAAG 1312
Db 254 SerArgAsnProIleTyrSerAsnAsnPheGlyLysPheGluIleThrProGlu--- 272
QY 1313 AAGAACCCCGAGCTTCAGGACCTGGACATGATGCTACCTGTGTAGAGATCAAAAGAGGA 1372
Db 273 LysAsnProGlnLeuArgAspLeuAspIlePheLeuSerSerValAspIleAsnGluGly 292
QY 1373 GCTTTGATGCTCCACATCTCAACTCAAGGCCATGTTATCGTCGTCGTCACAAAGGA 1432
Db 293 AlaLeuLeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGly 312
QY 1433 ACTGGAACCTTGAATCTCGTGTGAAGAAAGAGCAACACAGAGGGGACCGCGGAA 1492
Db 313 AspAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnLysGlnLysGlnGlu 331
QY 1493 GAAGAGGAGGACGAAGACGAAGAGAGGAGGGAAGTAACAGAGAGGTGCGTAGGTACACA 1552
Db 1552
```

```

Db 332 GluGluPro-----LeuGluValGlnArgTyrArg 341
QY 1553 GCGAGGTTGAAGAGGCGCATGTGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAAC 1612
Db 342 AlaGluLeuSerGluAspAspValPheValIleProAlaAlaTyrProPheValValAsn 361
QY 1613 GCTTCTCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTGAAACACCAACACCAATC 1672
Db 362 AlaThrSerAsnLeuAsnPheLeuAlaPheGlyIleAsnAlaGluAsnGlnArgAsn 391
QY 1673 TTCTTTCAGAGTGCATAAGACCAATGTGATAGACCAAGAGAGAGGAGGAGGATTTTA 1732
Db 382 PheLeuAlaGlyGluLysAspAsnValValArgGlnIleGluArgGlnValGlnGluLeu 401
QY 1733 GCATTCCTCCGTCGGTGGTGAACAAGTTGAGAGCTCATCAAAACACAGAGGAAATCTCAC 1792
Db 402 AlaPheProGlySerAlaGlnAspValGluArgLeuLeuLysLysGlnArgGluSerTyr 421
QY 1793 TTTGTGAGTGCCTCGTCTCTCAATCTCAATCTCGTCTCGTCTCGTCTCGTCTCGTCT 1852
Db 422 PheValAspAlaGlnProGln----- 428
QY 1853 CTTGAGAAAGAGGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1912
Db 429 -----GlnLysGluGluGlySerLysGlyArgLysGlyProPheProSerIle 444
QY 1913 TTGAAGGCT 1921
Db 445 LeuGlyAla 447

RESULT 10
US-10-245-227D-93
; Sequence 93, Application US/10245227D
; Patent No. 6936696
; GENERAL INFORMATION:
; APPLICANT: Monsanto Company
; APPLICANT: Rapp, William D.
; APPLICANT: Peng, JieXin
; APPLICANT: Nadig, Gautham
; APPLICANT: Venkatesh, Tyamagondlu
; TITLE OF INVENTION: Enhanced Protein and Methods of Their Use
; FILE REFERENCE: REN-00-087 US
; CURRENT APPLICATION NUMBER: US/10/245,227D
; CURRENT FILING DATE: 2002-09-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 93
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Glycine max
US-10-245-227D-93

Alignment Scores:
Pred. No.: 1,29e-104      Length: 449
Score: 1162.00           Matches: 236
Percent Similarity: 70.4%      Conservative: 90
Best Local Similarity: 51.0%    Mismatches: 83
Query Match: 32.4%           Indels: 54
DB: 2                     Gaps: 10

US-10-728-323-1 (1-2032) x US-10-245-227D-93 (1-449)

QY 542 GTGAGGAGAAACATCTCGGAACAAACCTTTTACTTC---CCGTCAAGCGGTTTAGC 598
Db 36 ValArgGluAsp-----GluAsnAsnProPheTyrLeuArgSerSerAsnSerPheGln 53
QY 599 ACCCGCTAGCGGAACCAAAACGCTAGGATCCGGTCTCTGAGAGGTTTTCACCAAGGTCA 658
Db 54 ThrLeuPheGluAsnGlnAsnGlyArgGileArgLeuGlnArgPheAsnLysArgSer 73
QY 659 AGGCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAACCTAACACT 718
Db 74 ProGlnLeuGluAsnLeuArgAspTyrArgIleValGlnPheGlnSerLysProAsnThr 93
QY 719 CTTGTTCTCCCAAGCAGCGTGTATGCTGATAACATCTTGTGTATCCAGCAGGCGAAGCC 778
Db 94 IleLeuLeuProHisHisAlaAspPheLeuLeuPheValLeuSerGlyArgAla 113
QY 779 ACCTGACCGTACCAATGCGCAATAACAGAAAGAGCTTTAATCTTCAGCAGGCGCATGCA 838
Db 114 IleLeuThrLeuValAsnAsnAspArgAspSerTyrAsnLeuHisProGlyAspAla 133
QY 839 CTCAGAAATCCATCCCGTTTCATTTCTCATCTTGAACCGCATGACACCAAGAACCTC 898
Db 134 GlnArgIleProAlaGlyThrTyrTyrLeuValAsnProHisAspHisGlnAsnLeu 153
QY 899 AGAGTAGCTAAATCTCCATGCCCGTTTAAACACACCCGCGCAGTTTGAGGATTTCTCCCG 958
Db 154 LysIleLeuLysLeuAlaIleProValAsnLysProGlyArgTyrAspAspPhePheLeu 173
QY 959 GCGAGCAGCGGACCAATCATCTCTTCGCGGCTTCAGAGGAAATACGTTTGGAGGCC 1018
Db 174 SerSerThrGlnAlaGlnSerTyrLeuGlnGlyPheSerHisAsnIleLeuGluThr 193
QY 1019 GCCTTCAATCGGAATTCATAGATACGAGGAGTGTCTGTAGAAGAGATCGAGGAGGT 1078
Db 194 SerPheHisSerGluPheGluGluIleAsnArgValLeuLeu-----Gly 208
QY 1079 GACCAAGAGAGAGGCGAGCGAGTGGAGTACTCGGAGTAGTGAGAACATGAAGGA 1138
Db 209 GluGluGluGluGlnArgGlnGln-----GluGly 218
QY 1139 GTGATAGTCAAGTGTCAAAGGACGCTTGAAGAACTTACTAAGCAGCGCTAAATCGCT 1198
Db 219 ValIleValGluLeuSerLysGluGlnIleArgGlnLeuSerArgArgAlaLysSerSer 238
QY 1199 TCAAGAAA-----GCTCCGAGAGAGGAGAGATATCAACCAACCAATCAACTTGAGA 1252
Db 239 SerArgLysThrIleSerSerGluAspGlu-----ProPheAsnLeuArg 253
QY 1253 GAAGGCGAGCCCATCTTCTTCAACATTTGGGAAGTTATTTCAGGTGAGCCAGACAAG 1312
Db 254 SerArgAsnProIleTyrSerAsnAsnPheGlyLysPheGluIleThrProGlu--- 272
QY 1313 AAGAACCCCGAGCTTCAGGACCTGGACATGATGCTACCTGTGTAGAGATCAAAAGAGGA 1372
Db 273 LysAsnProGlnLeuArgAspLeuAspIlePheLeuSerSerValAspIleAsnGluGly 292
QY 1373 GCTTTGATGCTCCACATCTCAACTCAAGGCCATGTTATCGTCGTCGTCACAAAGGA 1432
Db 293 AlaLeuLeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGly 312
QY 1433 ACTGGAACCTTGAATCTCGTGTGAAGAAAGAGCAACACAGAGGGGACCGCGGAA 1492
Db 313 AspAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnLysGlnLysGlnGlu 331
QY 1493 GAAGAGGAGGACGAAGACGAAGAGAGGAGGGAAGTAACAGAGAGGTGCGTAGGTACACA 1552
Db 1552
```

QY 719 CTTGTTCTCCCAAGCAGCGTGATGCTGATTAACATCTCTGTTATCCAGCAAGGCGCAAGCC 778
Db : : : : :
94 IleIleLeuProHisAlaAspPheIleIleIleLeuSerGlyArgAla 113
QY 779 ACCGTACCGTACCAATGCGCAATACAGAAAGAGCTTTAATCTTCCAGAGGCGCATGCA 938
Db : : : : :
114 IleLeuThrValAsnAsnAspArgAspSerTyrAsnIleHisProGlyAspAla 133
QY 839 CTCAGAAATCCATCCCGTTTCATTTCTCATCTTGAACCGCCATCACAACCAAGACCTC 898
Db : : : : :
134 GluArgIleProAlaGlyThrThrTyrTyrIleIleAsnProHisAspHisGlnAsnLeu 153
QY 899 AGATAGCTAAATCTCCATCCCGTTAAACACACCCCGCGAGTTTGAGGATTTCTTCCCG 958
Db : : : : :
154 LysIleIleLysIleAlaIleProValAsnLysProGlyArgTyrAspPheLeu 173
QY 959 GCGAGCAGCGGACCAATCATCTACTTGCAGGCTTCAGCAGGAATAGTTGGAGGCC 1018
Db : : : : :
174 SerSerThrGlnAlaGlnGlnSerTyrLeuGlnGlyPheSerHisAsnIleLeuGluThr 193
QY 1019 GCCTTCAATCGCAATTAATAGATACGAGGCTGCTGTAGAACAGAATGCAGGAGT 1078
Db : : : : :
194 SerPheHisSerGluPheGluGluIleAsnArgValLeuLeu-----Gly 208
QY 1079 GAGCAAGAGAGAGCGGAGCGGATCGAGTACTCGGAGTAGTGAGAACAAATGAAGA 1138
Db : : : : :
209 GluGluGluGlnArgGlnGln-----GluGly 218
QY 1139 GTGATAGTAAAGTGTAAAGGAGCAGCTTGAAGAACTTACTAAGCAGCTTAATCCGTC 1198
Db : : : : :
219 ValIleValGluLeuSerLysGluGlnIleArgGlnLeuSerArgArgAlaLysSerSer 238
QY 1199 TCAAGAAA-----GCCTCCGAGAGAGAGGAGATATCACCAACCAATCACTTGACA 1252
Db : : : : :
239 SerArgLysThrIleSerSerGluAspGlu-----ProPheAsnLeuArg 253
QY 1253 GAAGGCGAGCGCGATCTTTTAAACAATTTGGGAAGTTATTGAGTGAGGCGGAGCAAG 1312
Db : : : : :
254 SerArgAsnProIleTyrSerAsnAsnPheGlyLysPhePheGluIleThrProGlu--- 272
QY 1313 AGAAGCCCGAGCTCAGAGCCTGGACATGCTCCTCAGCTGTGTAGAGATCAAGAGGA 1372
Db : : : : :
273 LysAsnProGlnLeuArgAspLeuAspIlePheLeuSerSerValAspIleAsnGluGly 292
QY 1373 GCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCGTCTGCTCAACAAGGA 1432
Db : : : : :
293 AlaLeuLeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGly 312
QY 1433 ACTGGAACCTTGAACCTCGTGCTGTAAAGAAAGACAAACAGAGGGGAGCGCGGAA 1492
Db : : : : :
313 AspAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnGlnLysGlnLysGlnGlu 331
QY 1493 GAAGAGGAGACGAAGACGAGAGAGAGGAGGAAAGTAAACAGAGAGTGTGGTATACACA 1552
Db : : : : :
332 GluGluPro-----LeuGluValGlnArgTyrArg 341
QY 1553 GCGAGGTGAAGAGCGCGATGTTTCATCATGCCAGAGCTCATCCAGTAGCCATCAAC 1612
Db : : : : :
342 AlaGluLeuSerGluAspIleIleValIleProAlaIleTyrProPheValValAsn 361
QY 1613 GCTTCTCCGAACTCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1672
Db : : : : :
362 AlaThrSerAsnLeuAsnIleIleAlaIleGlyIleAsnAlaGluAsnAsnGlnArgAsn 381
QY 1673 TTCCTTGCAGGTATAGAGCAATGTATAGACCATGATAGAGAGCAAGCAAGGAGATTTA 1732
Db : : : : :
382 PheLeuAlaGlyLysAspAsnValValArgGlnIleGluArgGlnValGlnGluLeu 401
QY 1733 GCATTCCTCGGTCGGTGAACAAGTTGAGAGCTCATCAAAAACCAAGAGGAATCTCAC 1792
Db : : : : :
402 AlaPheProGlySerAlaGlnAspValGluArgLeuLeuLysLysGlnArgGluSerTyr 421

QY 1793 TTTGTGAGTGCTCGTCTCAATCTCAATCTCAATCTCGTCTCGTCTCGTCTCGTCT 1852
Db : : : : :
422 PheValAspAlaGlnProGln----- 428
QY 1853 CTTGAGAAAGAGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1912
Db : : : : :
429 -----GlnLysGluGluGlySerLysGlyArgLysGlyProPheProSerIle 444
QY 1913 TTTGAAGGCT 1921
Db : : : : :
445 LeuGlyAla 447
RESULT 11
US-10-245-227D-14
; Sequence 14, Application US/10245227D
; Patent No. 6936696
; GENERAL INFORMATION:
; APPLICANT: Monsanto Company
; APPLICANT: Rapp, William D.
; APPLICANT: Peng, Jixin
; APPLICANT: Nadig, Gautham
; APPLICANT: Venkatesh, Tyamagondlu
; TITLE OF INVENTION: Enhanced Protein and Methods of Their Use
; FILE REFERENCE: REN-00-087 US
; CURRENT APPLICATION NUMBER: US/10/245,227D
; CURRENT FILING DATE: 2002-09-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 14
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: mature form of beta-conglycinin beta-subunit
US-10-245-227D-14
Alignment Scores:
Pred. No.: 3,786-104 Length: 417
Score: 1157.00 Matches: 237
Percent Similarity: 70.0% Conservative: 87
Best Local Similarity: 51.2% Mismatches: 85
Query Match: 32.3% Indels: 54
DB: 2 Gaps: 10
US-10-728-323-1 (1-2032) x US-10-245-227D-14 (1-417)
QY 542 GTGAGGAGAAACATCTCGGAGCAACCTTTCTACTTC---CCGTCAAGCGGTTTACG 598
Db 4 ValArgGluAsp-----GluAsnAsnProPheTyrLeuArgSerSerAsnSerPheGln 21
QY 599 ACCCGCTACGGAAACCAAAACCGTAGGATCCGGGTCCTCGAGAGTTTGCACCAAGGTCA 658
Db : : : : :
22 ThrLeuPheGluAsnGlnAsnGlyArgIleArgLeuLeuGlnArgPheAsnLysArgSer 41
QY 659 AGCGAGTTTCAAGATCTCCAGAAATCACCGTATTGTGAGATCGAGGCCAAACCTAACACT 718
Db : : : : :
42 ProGlnLeuGluAsnLeuArgAspTyrArgIleValGlnPheGlnSerLysProAsnThr 61
QY 719 CTTGTTCTCCCAAGCAGCGTGATGCTGATTAACATCTCTGTTATCCAGCAAGGCGCATG 778
Db : : : : :
62 IleLeuLeuProHisAlaAspPheLeuLeuPheValLeuSerGlyArgAla 81
QY 779 ACCGTGACGTAGCAATGGCAATTAACAGAAAGAGCTTTAAATCTTGACGAGGCCATGCA 838
Db : : : : :
82 IleLeuThrLeuValAsnAsnAspArgAspSerTyrAsnLeuHisProGlyAspAla 101
QY 839 CTCAGAAATCCATCCCGTTTCATTTCTTCAATCTTGAACCGCCATGACCAACGAGACCTC 898
Db : : : : :
102 GlnArgIleProAlaGlyThrThrTyrTyrLeuValAsnProHisAspHisGlnAsnLeu 121
QY 899 AGATAGTAAATCTCCATCCCGTTTAAACACACCCCGCGAGTTTGAGGATTTCTTCCCG 958
Db : : : : :
122 LysIleIleLysLeuAlaIleProValAsnLysProGlyArgTyrAspAspPhePheLeu 141

```
QY 959 GCGAGAGCGGAGACCAATCATCTTCTGACGGGCTTCAGCAGGAATACGTTGGAGGCC 1018
Db 142 SerSerThrGlnAlaGlnGlnSerTyrLeuGlnGlnPheSerHisAsnLeuLeuGluThr 161
QY 1019 GCCTTCAATCGGAAATTCAATGAGATACGAGGGTCTGTGTAGAACGAGAATGCAGGAGGT 1078
Db 162 SerPheHisSerGluPheGluGluIleAsnArgValLeuLeu-----Gly 176
QY 1079 GAGCAAGAGAGAGAGCGGAGCGATCGAGTACTCGGAGTAGTGAGAACCAATGAAGGA 1138
Db 177 GluGluGluGluGlnArgGlnGln-----GluGly 186
QY 1139 GTGATAGTCAAACTGTCAAAGGAGCAGCTGTGAAGAACTTACTTAAGCAGCGTAAATCCGTC 1198
Db 187 ValIleValGluLeuSerLysGluGlnIleArgGlnLeuSerArgAlaLysSerSer 206
QY 1199 TCAAGAGAA-----GCCTCCGAGAGAGAGGAGATATCACCAACCCCAATCAACTTGAGA 1252
Db 207 SerArgLysThrIleSerSerGluAspGlu-----ProPheAsnLeuArg 221
QY 1253 GAAGGGGAGCGCCGATCTTTCTAAACAATTGGGAAGTTATTGAGGTGAAGCCAGCAGACAAG 1312
Db 222 SerArgAsnProIleTyrSerAsnAsnPheGlyLysPheGluIleThrProGlu--- 240
QY 1313 AAGAACCCCGAGCTTCAGGACCTGGACATGATCTCACCTGTGTAGATCAAGAAGGA 1372
Db 241 LysThrProGlnLeuArgAspLeuAspIlePheLeuSerSerValAspIleAsnGluGly 260
QY 1373 GCTTTGATGCTCCACACTTCAACTCAAGGCCATGTTATCTGTCGTGTCACAAAGGA 1432
Db 261 AlaLeuLeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGly 280
QY 1433 ACTGGAACCTTGAACCTCGTGGCTGTAAGAAAAAGCAACAACAGAGGGGACGGCGGAA 1492
Db 281 AspAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnGlnLysGlnLysGlnGlu 299
QY 1493 GAAGAGGAGGACCAAGCAGNAGAGAGGGGAAGTAACAGAGAGGTGCGTAGTACACA 1552
Db 300 GluGluPro-----LeuGluValGlnArgTyrArg 309
QY 1553 GCGAGGTGAAGGAGCGGATGTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAAC 1612
Db 310 AlaGluLeuSerGluAspValPheValIleProAlaAlaTyrProPheValValAsn 329
QY 1613 GCTTCCTCCGAATCCATCTGCTTGGCTTCGGTATCAACGCTCAAAACCAACCAAGATC 1672
Db 330 AlaThrSerAsnLeuAsnPheLeuAlaPheGlyIleAsnAlaGluAsnGlnArgAsn 349
QY 1673 TTCCTTGACGATTAAGNCAATGTGATAGACAGATAGAGAACGACGAGGATTTA 1732
Db 350 PheLeuAlaGlyLysAspAsnValValArgGlnIleGluArgGlnValGlnGluLeu 369
QY 1733 GCATTCCCTCGGTGGGTGAACAAGTTGAGAAGCTCATCAAAAACCAAGAGGAATCTCAC 1792
Db 370 AlaPheProGlySerAlaGlnAspValGluArgLeuLeuLysLysGlnArgGluSerTyr 389
QY 1793 TTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCCGTCGTCTCCTGAGAAGAGTCT 1852
Db 390 PheValAspAlaGlnProGln----- 396
QY 1853 CCTGAGNAGAGATCAAGGAGGAGAAACCAAGGAGGGAGGTCCTCTTTCAATT 1912
Db 397 -----GlnLysGluGluGlySerLysGlyArgLysGlyProPheProSerIle 412
QY 1913 TTGAAGGCT 1921
Db 413 LeuGlyAla 415
```

RESULT 12

US-10-245-227D-12

; Sequence 12, Application US/10245227D

; Patent No. 6936696

```
; GENERAL INFORMATION:
; APPLICANT: Monsanto Company
; APPLICANT: Rapp, William D.
; APPLICANT: Peng, Jiexin
; APPLICANT: Nadiq, Gautham
; APPLICANT: Venkatesh, Tyamagondlu
; TITLE OF INVENTION: Enhanced Protein and Methods of Their Use
; FILE REFERENCE: REN-00-087 US
; CURRENT APPLICATION NUMBER: US/10/245,227D
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 12
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: epitope form of beta-conglycinin beta subunit
US-10-245-227D-12
```

```
Alignment Scores:
Pred. No.: 3.83e-104 Length: 425
Score: 1157.00 Matches: 237
Percent Similarity: 70.0% Conservative: 87
Best Local Similarity: 51.2% Mismatches: 85
Query Match: 32.3% Indels: 54
DB: 2 Gaps: 10
```

US-10-728-323-1 (1-2032) x US-10-245-227D-12 (1-425)

```
QY 542 GTGAGGAGAGAAACATCTCGGAACACCTTCTACTTC--CCGTCAAGCGGTTTAGC 598
Db 12 ValArgGluAsp-----GluAsnAsnProPheTyrLeuArgSerSerAsnSerPheGln 29
QY 599 ACCCGCTAGCGGAACCAAAACGCTAGGATCCGGGTCCTCGCAGAGGTTTGACCAAGGTCA 658
Db 30 ThrLeuPheGluAsnGlnAsnGlyArgIleArgLeuLeuGlnArgPheAsnLysArgSer 49
QY 659 AGCAGTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGGCCAACCTAACACT 718
Db 50 ProGlnLeuGluAsnLeuArgAspTyrArgIleValGlnPheGlnSerLysProAsnThr 69
QY 719 CTGTGTTCTTCCCAAGCACCGCTGATGCTATAACATCCTTGTATTCCAGCAAGGCGAAGCC 778
Db 70 IleLeuLeuProHisIleAlaAspAlaAspPheLeuLeuPheValLeuSerGlyArgAla 89
QY 779 ACGTGACCGTAGCAAAATGGCAATAACAGAAAAGAGCTTTAATCTTCAGGAGGCCATGCA 838
Db 90 IleLeuThrLeuValAsnAsnAspAspArgAspSerTyrAsnLeuHisProGlyAspAla 109
QY 839 CTCAGAAATCCATCCCGTTTCATTCTCTACATCTTGAAACGCGCATCACACCAAGACCTC 898
Db 110 GlnArgIleProAlaGlyThrThrTyrTyrLeuValAsnProHisAspHisGlnAsnLeu 129
QY 899 AGAGTAGCTAAATCTCCATGCCCGTTAAACACACCGCGCAGTTTGAGGAGTTTCTTCCCG 958
Db 130 LysIleIleLeuLeuAlaIleProValAsnLysProGlyArgTyrAspAspPhePheLeu 149
QY 959 GCGAGCAGCGAGACCAATCATCTACTTTCAGGGGCTTCAGCAGGAATACGTTGGAGGCC 1018
Db 150 SerSerThrGlnAlaGlnGlnSerTyrLeuGlnGlyPheSerHisAsnIleLeuGluThr 169
QY 1019 GCCTTCAATCGGAATTCATGATACGAGGAGTCTGTGTAGAACGAGAATGCAGGAGGT 1078
Db 170 SerPheHisSerGluPheGluGluIleAsnArgValLeuLeu-----Gly 184
QY 1079 GAGCAAGAGAGAGAGCGGAGGCGATCGAGTACTCGGAGTAGTGAGAACCAATGAAGGA 1138
Db 185 GluGluGluGluGlnArgGlnGln-----GluGly 194
QY 1139 GTGATAGTCAAACTGTCAAAGGAGCAGCTGTGAAGAACTTACTTAAGCAGCGTAAATCCGTC 1198
Db 195 ValIleValGluLeuSerLysGluGlnIleArgGlnLeuSerArgAlaLysSerSer 214
```

```
QY 1199 TCAGAGAA-----GGCTCCGAAGAGAGGAGATATCACCAACCCCAATCACTTGAGA 1252
Db 215 SerArgLysThrIleSerSerGluAspGlu-----ProPheAsnLeuArg 229
QY 1253 GAAGGAGAGCCGATCTTCTTAACACTTTGGGAAGTATTGAGGTGAAGCAGACAAAG 1312
Db 230 SerArgAsnProIleTyrSerAsnAsnPheGlyLysPheGluIleThrProGlu--- 248
QY 1313 AAGAACCCCGAGCTCAGACCTGAGCATGATGCTCACCTGCTGAGATCAAGAAAGA 1372
Db 249 LysThrProGlnLeuArgAspLeuAspIlePheLeuSerSerValAspIleAsnGluGly 268
QY 1373 GCTTTGATGCTCCACACTTCAACTCAAGGCGCATGTTATCGTCTGCTCAACAAGGA 1432
Db 269 AlaLeuLeuLeuProHisPheAsnSerLysAlaIleValIleuValIleAsnGluGly 288
QY 1433 ACTGGAACCTTGAACTCGTGTGTAAGAAAGAGCAACAACAGAGGGGACGGCGGAA 1492
Db 289 AspAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnGlnLysGlnGlu 307
QY 1493 GAAGAGGAGCGAAGACGAAGAGGAGGAGTAAACAGAGAGGTGCTAGGTACACA 1552
Db 308 GluGluPro-----LeuGluValGlnArgTyrArg 317
QY 1553 GCGAGGTTGAAGAGCGATGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAAC 1612
Db 318 AlaGluLeuSerGluAspValPheValIleProAlaIleTyrProPheValValAsn 337
QY 1613 GCTTCTCCGAATCCTCATCTGCTTGGCTTCGGTATCAACGCTGAAACCAACACAGATC 1672
Db 338 AlaThrSerAsnLeuAsnPheLeuAlaPheGlyIleAsnAlaGluAsnAsnGlnArgAsn 357
QY 1673 TTCTTGAGGTGATAGCAATGATGATAGCAGATAGACAGACGAGGAGGATTTA 1732
Db 358 PheLeuAlaGlyGluLysAspAsnValValArgGlnIleGluArgGlnValGlnGluLeu 377
QY 1733 GCATTCCTCGGTGGGTGAAACAAAGTTGAGAAAGCTCATCAAAACCAAGAGGAATCTCAC 1792
Db 378 AlaPheProGlySerAlaGlnAspValGluArgLeuLeuLysGlnArgGluSerTyr 397
QY 1793 TTTGTGAGTGTCTCTCAATCTCAATCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCT 1852
Db 398 PheValAspAlaGlnProGln-----404
QY 1853 CTGAGAAAGAGATCAAGAGGAGGAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1912
Db 405 -----GlnLysGluGluGlySerLysGlyArgLysGlyProPheProSerIle 420
QY 1913 TTGAAGGCT 1921
Db 421 LeuGlyAla 423
```

RESULT 13
US-10-245-227D-2
; Sequence 2, Application US/10245227D
; Patent No. 6936696
; GENERAL INFORMATION:
; APPLICANT: Monsanto Company
; APPLICANT: Rapp, William D.
; APPLICANT: Peng, Jixian
; APPLICANT: Nadiq, Gautham
; APPLICANT: Venkatesh, Tymanagondlu
; TITLE OF INVENTION: Enhanced Protein and Methods of Their Use
; FILE REFERENCE: REN-00-087 US
; CURRENT APPLICATION NUMBER: US/10/245, 227D
; CURRENT FILING DATE: 2002-09-17
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Glycine max

US-10-245-227D-2

Alignment Scores:
Pred. No.: 2, 67e-100 Length: 390
Score: 1117.50 Matches: 225
Percent Similarity: 72.6% Conservative: 83
Best Local Similarity: 53.1% Mismatches: 79
Query Match: 31.2% Indels: 37
DB: 2 Gaps: 9

US-10-728-323-1 (1-2032) x US-10-245-227D-2 (1-390)

```
QY 542 GTGAGGAGAGAAACATCTCGGAACAAACCTTTCTACTTC---CCGTCAGGCGGTTTACG 598
Db 1 ValArgGluAsp-----GluAsnAsnProPheTyrPheArgSerSerAsnSerPheGln 18
QY 599 ACCCGCTACGGGACCAAAACGGTAGGATCCGGGTCTCGCAGAGGTTTACCAAAAGGTCA 658
Db 19 ThrLeuPheGluAsnGlnAsnValArgIleArgLeuLeuGlnArgPheAsnLysArgSer 38
QY 659 AGGCAGTTTCAAGATCTCCAGATCACCGTATTGTGCAGATCGAGGCCAAACCTAACACT 718
Db 39 ProGlnLeuGluAsnLeuArgAspTyrArgIleValGlnPheGlnSerLysProAsnThr 58
QY 719 CTTGTTCTCCACACGACGCTGATGCTGATACATCTCTTTATCCAGCAAGGGCAAGCC 778
Db 59 IleLeuLeuProHisHisAlaAspAlaAspPheLeuLeuPheValLeuSerGlyArgAla 78
QY 779 ACCGTGACGTAGCAAAATGGCAATTAACAGAAAGAGCTTTAATCTTGCAGGGGCCATGCA 838
Db 79 IleLeuThrLeuValAsnAsnAspArgAspSerTyrAsnLeuHisProGlyAspAla 98
QY 839 CTCGAATCCCATCCCGTTTCATTTCTTACATCTTGAACCGCCATCAACCAACAGAACCTC 898
Db 99 GlnArgIleProAlaGlyThrThrTyrTyrLeuValAsnProHisAspHisGlnAsnLeu 118
QY 899 AGAGTAGTAAATCTCCATGCCCTTAACACACCCGCCAGCTTTGAGGATTTCTTCCCG 958
Db 119 LysIleIleLysLeuAlaIleProValAsnLysProGlyArgTyrAspAspPhePheLeu 138
QY 959 GCGACGCGGAGACCAATCATCTTCTGAGGCTTTCAGGAGGAGGAGGAGGAGGAGGAGG 1018
Db 139 SerSerThrGlnAlaGlnGlnSerTyrLeuGlnGlyPheSerHisAsnIleLeuGluThr 158
QY 1019 GCCTTCAATCGGAATTCGAATAGATACGAGGGTCTGTTAGAGAAGAGAGATCGAGAGT 1078
Db 159 SerPheHisSerGluPheGluGluIleAsnArgValLeuPhe-----Gly 173
QY 1079 GAGCAAGAGGAGAGAGGCGAGGCGATGGAGTACTCGGAGTAGTGAGACAATGAAGGA 1138
Db 174 GluGluGluGluGlnArgGlnGln-----GluGly 183
QY 1139 GTGATAGTCAAGTGTCAAGGAGCAGCGTTGAAGAACTTACTTAAGCACGCTAAATCCGTC 1198
Db 184 ValIleValGluLeuSerLysGluGlnIleArgGlnLeuSerArgArgAlaLysSerSer 203
QY 1199 TCAAAAGAAA-----GGCTCCGAAGAGAGGAGATATCACCAACCCCAATCACTTGAGA 1252
Db 204 SerArgLysThrIleSerSerGluAspGlu-----ProPheAsnLeuArg 218
QY 1253 GAAGGCGAGCCGATCTTTCTTAACACTTTGGGAAGTATTGAGGTGAAGCAGACAAAG 1312
Db 219 SerArgAsnProIleTyrSerAsnAsnPheGlyLysPhePheGluIleThrProGlu--- 237
QY 1313 AAGAACCCCGAGCTTCAAGCCTGAGCATGATGCTCACCTGCTGAGATCAAGAAAGA 1372
Db 238 LysAsnProGlnLeuArgAspLeuAspIlePheLeuSerSerValAspIleAsnGluGly 257
QY 1373 GCTTTGATGCTCCACACTTCAACTCAAGGCGCATGTTATCGTCTGCTCAACAAGGA 1432
Db 258 AlaLeuLeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGly 277
QY 1433 ACTGAAACCTTTGAATCTCGTGGCTGTAGAAAGAGCAACAACAGAGGGGACGGCGGAA 1492
```


GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 23:41:32 ; Search time 6.9905 Seconds
(without alignments)
4094.130 Million cell updates/sec

Title: US-10-728-323-1

Perfect score: 3586

Sequence: 1 aataatcatatatattcatc.....cgttgtgctgtttcttccc 2032

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 250354 seqs, 46948837 residues

Total number of hits satisfying chosen parameters: 500708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/abs/ABSSWEB.spool/US10728323/runat_15052006.172148.22650/app.query.fasta.1
-DB=Published Applications_AA_New -QFMT=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs02p
-USER=US10728323 @CGN 1.1.63 @runat_15052006.172148.22650 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_AA_New.*

1: /SID55/ptodata/1/pubpaa/US08_NEW_PUB.pep.1.*
2: /SID55/ptodata/1/pubpaa/US06_NEW_PUB.pep.1.*
3: /SID55/ptodata/1/pubpaa/US07_NEW_PUB.pep.1.*
4: /SID55/ptodata/1/pubpaa/US08_NEW_PUB.pep.1.*
5: /SID55/ptodata/1/pubpaa/PCT_NEW_PUB.pep.1.*
6: /SID55/ptodata/1/pubpaa/US09_NEW_PUB.pep.1.*
7: /SID55/ptodata/1/pubpaa/US09_NEW_PUB.pep.1.*
8: /SID55/ptodata/1/pubpaa/US10_NEW_PUB.pep.1.*
9: /SID55/ptodata/1/pubpaa/US10_NEW_PUB.pep.1.*
10: /SID55/ptodata/1/pubpaa/US11_NEW_PUB.pep.1.*
11: /SID55/ptodata/1/pubpaa/US11_NEW_PUB.pep.1.*
12: /SID55/ptodata/1/pubpaa/US60_NEW_PUB.pep.1.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3286	91.6	626	9	US-10-498-026-89 Sequence 89, Appl
2	3286	91.6	626	11	US-11-033-039-10 Sequence 10, Appl
3	228.5	6.4	499	9	US-10-508-263-94 Sequence 94, Appl
4	179	5.0	8746	11	US-11-098-686-10232 Sequence 10232, A
5	172	4.8	798	11	US-11-072-512-3335 Sequence 3335, Ap

6	170.5	4.8	507	11	US-11-033-039-96 Sequence 96, Appl
7	166.5	4.6	481	9	US-10-508-263-28 Sequence 28, Appl
8	163.5	4.6	651	11	US-11-128-660-1 Sequence 1, Appl
9	162.5	4.5	499	11	US-11-087-099-1159 Sequence 1159, Ap
10	162	4.5	5689	9	US-10-784-004-709 Sequence 709, Appl
11	157	4.4	495	9	US-10-508-263-20 Sequence 20, Appl
12	153	4.3	677	9	US-10-982-545-12 Sequence 12, Appl
13	149	4.2	623	11	US-11-072-512-3378 Sequence 3378, Ap
14	148	4.1	562	9	US-10-508-263-24 Sequence 24, Appl
15	147.5	4.1	715	11	US-11-072-512-3385 Sequence 3385, Ap
16	146	4.1	1719	11	US-11-234-786-378 Sequence 378, Appl
17	145.5	4.1	656	9	US-10-784-004-451 Sequence 451, Appl
18	145	4.0	1197	11	US-11-087-099-8238 Sequence 8238, Ap
19	144	4.0	1212	9	US-10-979-095-9 Sequence 9, Appl
20	143	4.0	671	11	US-11-234-786-380 Sequence 380, Appl
21	143	4.0	671	11	US-11-139-041-306 Sequence 306, Appl
22	142.5	4.0	656	11	US-11-234-786-379 Sequence 379, Appl
23	142.5	4.0	656	11	US-11-139-041-305 Sequence 305, Appl
24	140.5	3.9	2910	9	US-10-330-773-39 Sequence 39, Appl
25	139	3.9	2004	9	US-10-469-469-250 Sequence 250, Appl
26	139	3.9	4373	10	US-11-118-524-2 Sequence 2, Appl
27	139	3.9	4374	11	US-11-128-572-2 Sequence 2, Appl
28	136.5	3.8	2426	11	US-11-203-806A-11 Sequence 11, Appl
29	135.5	3.8	2472	11	US-11-203-806A-2 Sequence 2, Appl
30	134	3.7	585	9	US-10-784-004-712 Sequence 712, Appl
31	134	3.7	795	9	US-10-770-726-49 Sequence 49, Appl
32	133.5	3.7	594	9	US-10-131-826A-10 Sequence 10, Appl
33	133.5	3.7	594	9	US-10-973-115B-10 Sequence 10, Appl
34	133.5	3.7	594	9	US-10-137-873A-10 Sequence 10, Appl
35	133.5	3.7	594	9	US-10-152-370-10 Sequence 10, Appl
36	133.5	3.7	594	11	US-11-290-153-10 Sequence 10, Appl
37	132.5	3.7	676	11	US-11-135-855-28 Sequence 28, Appl
38	132.5	3.7	717	11	US-11-135-855-29 Sequence 29, Appl
39	132.5	3.7	841	11	US-11-294-997-37 Sequence 37, Appl
40	131.5	3.7	2344	9	US-10-330-773-627 Sequence 627, Appl
41	131	3.7	671	8	US-10-196-749-346 Sequence 346, Appl
42	131	3.7	671	9	US-10-194-487-346 Sequence 346, Appl
43	131	3.7	671	9	US-10-195-883-346 Sequence 346, Appl
44	131	3.7	671	9	US-10-195-888-346 Sequence 346, Appl
45	131	3.7	671	9	US-10-195-889-346 Sequence 346, Appl

ALIGNMENTS

RESULT 1
US-10-498-026-89
; Sequence 89, Application US/10498026
; Publication No. US20060024334A1
; GENERAL INFORMATION:
; APPLICANT: CIRCASSIA LIMITED
; TITLE OF INVENTION: IMMUNOTHERAPEUTIC METHODS AND SYSTEMS
; FILE REFERENCE: N.87430 WO GCW
; CURRENT APPLICATION NUMBER: US/10/498,026
; CURRENT FILING DATE: 2004-06-04
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 89
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-498-026-89

Alignment Scores:
Pred. No.: 3,43e-262 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.6% Indels: 0
DB: 9 Gaps: 0

US-10-728-323-1 (1-2032) x US-10-498-026-89 (1-626)

QY 50 ATGAGAGGAGGGTTTCTCCACTGATGCTGTGCTAGGATCCTTGCTCTGCTTCAGTT 109

Db 1 MetArgGlyArgValSerProLeuMetLeuLeuGlyLeuValLeuAlaSerVal 20
QY 110 TGTGCAACGATGCCCAAGTCATCACCTTACCAAGAAACAGAGAACCCCTGCGCCAG 169
Db 21 SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln 40
QY 170 AGTGTCTCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGATCGAGTCT 229
Db 41 ArgCysLeuGlnSerCysGlnGlnProAspAspLeuLysGlnLysAlaCysGluSer 60
QY 230 CGCTGCACCAAGCTCCAGATATGATCCTCGTTGTCTATGATCCTCGAGACACACTGCC 289
Db 61 ArgCysThrLysLeuGlnLysArgProArgProArgCysValTyrAspProArgGlyHisThrGly 80
QY 290 ACCACCAACCAACGTTCCCTCCAGGGGAGCGGACACGCTGGCGGCCCAACCCGGAGACTAC 349
Db 81 ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr 100
QY 350 GATGATGACCGCGCTCAACCCGGAAGAGAGGAAGGCGCGATGGGACACAGCTGCACCG 409
Db 101 AspAspAspArgArgGlnProArgArgGluGluGlyGlyArgTrpGlyProAlaGlyPro 120
QY 410 AGGAGCGCTCAAGAGCAAGAACTCGAGACCAACCAAGAGAACTTGGAGGCGACCAAGT 469
Db 121 ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer 140
QY 470 CATCAGCAGCCACGGAATAAGGCCCGAAGGAGAGAGGAACAAGAGTGGGAACA 529
Db 141 HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGluTrpGlyThr 160
QY 530 CAGGTTAGCCATGTGAGGGAAGAAACATCTCGGAACAACCCCTTCTACTTCCGCTCAAG 589
Db 161 ProGlySerHisValArgGluThrSerArgAsnAsnProPheTyrPheProSerArg 180
QY 590 CGGTTTAGCACCGCTACGGGAACCAACCGGTAGGATCCGGTCTCGCAGAGGTTTGAC 649
Db 181 ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp 200
QY 650 CAAAGTCAAGGCGATTTCCAGAACTCCAGAAATCCACCGTATTGTGCAGATCGAGGCCAAA 709
Db 201 GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGluAlaLys 220
QY 710 CCTAACACTTGTCTTCCAGACGCGTGATGATGATTAACATCTTGTATTCACGAA 769
Db 221 ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln 240
QY 770 GGGCAAGCCACCGTAGCCGTAAGCAATGGCAATAACAGAAAGAGCTTTAATCTTGACGAG 829
Db 241 GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu 260
QY 830 GGCATCACTCAAGATCCCATCCGTTTCATTTCTTACATCTTGAAACCGCCATGACAAC 889
Db 261 GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn 280
QY 890 CAGAACTCAGACTAGCTAAATCTCCATCCCGCTTAACACACCCGCGCAGTTTGAGGAT 949
Db 281 GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp 300
QY 950 TTCTTCCGCGAGCAGCCGAGACCAATCATCTTCTTCTCAGGCGCTTCAGCAGGAATACG 1009
Db 301 PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr 320
QY 1010 TTGAGGCGCCCTTCAATCGGAATTCATATGAGATACGAGGAGTGTCTGTAGAGAGAAT 1069
Db 321 LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn 340
QY 1070 GCAGGAGCTCAGCAAGGAGAGAGGGCAGAGCGGATGGAGTACTCGGACTACTGAGAAC 1129
Db 341 AlaGlyGlyGluGlnGluGluArgGlyGlnArgTrpSerThrArgSerSerGluAsn 360
QY 1130 AATGAAGGAGTGTAGTCAAAAGTGTCAAAGGAGCAGCTTGAAGAACTTACTAAGCAGCT 1189

Db 361 AsnGluGlyValIleValLysValSerLysGluHisValGluLeuThrLysHisAla 380
QY 1190 AAATCGCTCTCAAGAAAGGCTCCGAAGAGAGGAGATATCACCAACCAATCACTTG 1249
Db 381 LysSerValSerLysLysGlySerGluGluGlyAspIleThrAsnProIleAsnLeu 400
QY 1250 ACAGAAAGCGGACCCGATCTTCTTAACAACCTTTGGGAAGTATTATTGAGGTGAAGCCAGAC 1309
Db 401 ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp 420
QY 1310 AAGAAGAAACCCCGACTTCAGGACCTTGGACATGATCTCACCTGTGTAGAGATCAAGAA 1369
Db 421 LysLysAsnProGlnLeuGlnAspLeuAspMetLeuThrCysValGluIleLysGlu 440
QY 1370 GGAGCTTGTATGCTCCACACTTCAACTCAAAAGGCCATGGTTATCGTCTGCTCAACAA 1429
Db 441 GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys 460
QY 1430 GGAACCTGGAACCTTGAACCTCGTGTAAAGAAAAGAGCAACAACAGAGGGGACGCGG 1489
Db 461 GlyThrGlyAsnLeuGluLeuValAlaValArgLysGluGlnGlnGlnArgGlyArgArg 480
QY 1490 GAAGAAGAGGAGGACGAAAGAGAGAGGAGGAAAGTAAACAGAGAGTGCCTAGGTAC 1549
Db 481 GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr 500
QY 1550 ACAGCGAGTTCAAGGAAGCGCATGTTCATCATGCCAGCAGCTCATCCAGTAGCCATC 1609
Db 501 ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle 520
QY 1610 AAGCTTCTCTCGAACTCCATCTGCTTGGCTTCGGTATCAACGCTCAAAACCAACACAGA 1669
Db 521 AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg 540
QY 1670 ATCTTCTCTGAGGTGATAAGGACAAATGTATAGACCAATGTATAGACAGAGAGCAAGAGGAT 1729
Db 541 IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp 560
QY 1730 TTAGCAATTCCTGGGTGGGTGAACAAGTTGAGAAGCTCATCAAAACCAAGAGGATCT 1789
Db 561 LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer 580
QY 1790 CACTTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCCGCTGCTCTCTGAGAAAGAG 1849
Db 581 HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu 600
QY 1850 TCTCTGAGAAAGAGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1909
Db 601 SerProGluLysGluAspGlnGluGluGluAsnGlnGlyGlyLysGlyProLeuLeuSer 620
QY 1910 ATTTTGAAGCTTTTAAAC 1927
Db 621 IleLeuLysAlaPheAsn 626

RESULT 2
US-11-033-039-10
; Sequence 10, Application US/11033039
; Publication No. US20060002947A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: LI-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2017US01
; CURRENT APPLICATION NUMBER: US/11/033,039
; PRIOR FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: 10/245,871
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 1452
; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 10
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-11-033-039-10

Alignment Scores:

Pred. No.: 3,43e-262 Length: 626
Score: 3286.00 Matches: 626
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 91.6% Indels: 0
DB: 11 Gaps: 0

US-10-728-323-1 (1-2032) x US-11-033-039-10 (1-626)

QY	50	ATGAGAGGAGGGTTTCTCCACTGATGCTTGCTAGGGATCCTTGCTCGCTTCAGTT	109
DB	1	MetArgGlyArgValSerProLeuMetLeuLeuLeuGlyIleLeuValLeuAlaSerVal	20
QY	110	TCTGCAACGCATGCCAAGTCATCACCTTACAGAGAGAAACAGAGAACCCCTGCGCCAG	169
DB	21	SerAlaThrHisAlaLysSerSerProTyrGlnLysLysThrGluAsnProCysAlaGln	40
QY	170	AGGTGCTCCAGAGTTGTCAACAGGAACCGGATGACTTGAAGCAAAAGGCATGCGAGTCT	229
DB	41	ArgCysLeuGlnSerCysGlnGlnGluProAspAspLeuLysGlnLysAlaCysGluSer	60
QY	230	CGCTGCACCAAGCTCGAGTATGATCTCTGTTGTGCTATGATCTTCGAGGACACACTGGC	289
DB	61	ArgCysThrLysLeuGluTyrAspProArgCysValTyrAspProArgGlyHisThrGly	80
QY	290	ACCACCAACCAAGTTTCCCTCCAGGGAGCGGACACGTTGGCGGCCCAACCCGGAGACTAC	349
DB	81	ThrThrAsnGlnArgSerProProGlyGluArgThrArgGlyArgGlnProGlyAspTyr	100
QY	350	GATGATACCGCGCTCAACCCCGAAGAGAGAGAGGCGCGGATGGGACCACTGGACCG	409
DB	101	AspAspAspArgArgGlnProArgGluGluGlyArgTyrGlyProAlaGlyPro	120
QY	410	AGGGAGCGTGAAGAGAGAAAGACTGGGACACCAACNAGAGAGATTGGAGGCGACCAAGT	469
DB	121	ArgGluArgGluArgGluAspTrpArgGlnProArgGluAspTrpArgArgProSer	140
QY	470	CATCAGCACCCAGGAAATAAGGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	529
DB	141	HisGlnGlnProArgLysIleArgProGluGlyArgGluGlyGluGlnGlnIleArgGlyThr	160
QY	530	CCAGGTAGCCATGTGAGGAGAAACATCTCGGAACAAACCTTTCTACTTCCCGTCAAGG	589
DB	161	ProGlySerHisValArgGluGluThrSerArgAsnAsnProPheTyrPheProSerArg	180
QY	590	CGTTTATGACCGCTACGGGAAACCAACCGTAGATCCGGTCTTCGAGAGTTTGAC	649
DB	181	ArgPheSerThrArgTyrGlyAsnGlnAsnGlyArgIleArgValLeuGlnArgPheAsp	200
QY	650	CAAAAGTCAAGCAGTTTCAGAAATCTCCAGAAATCCCGTATTGTGCAGATCCAGAGCCCAA	709
DB	201	GlnArgSerArgGlnPheGlnAsnLeuGlnAsnHisArgIleValGlnIleGlnAlaLys	220
QY	710	CCTAACACTTTGTTCTTCCCAAGCACGCTGATGCTGATAACATCTTGTATCCAGCAA	769
DB	221	ProAsnThrLeuValLeuProLysHisAlaAspAlaAspAsnIleLeuValIleGlnGln	240
QY	770	GGGCAAGCCACGCTGACCGTAGCAATGGCAATACAGAAAGAGCTTTAATCTTGACGAG	829
DB	241	GlyGlnAlaThrValThrValAlaAsnGlyAsnAsnArgLysSerPheAsnLeuAspGlu	260
QY	830	GGCCATGCATCAGAAATCCATCCGGTTTCATTTCTCATCTTCTGAACCGCCATCACAC	889
DB	261	GlyHisAlaLeuArgIleProSerGlyPheIleSerTyrIleLeuAsnArgHisAspAsn	280
QY	890	CAGAACCTCAGATAGTAAAAATCTCCATGCCCGTTAAACACACCCCGCCAGTTTGAGGAT	949

DB	281	GlnAsnLeuArgValAlaLysIleSerMetProValAsnThrProGlyGlnPheGluAsp	300
QY	950	TTTCTCCCGCGAGCAGCGAGACCAATCATCTACTTGCAGCGCTTCAGCAGGAATAGC	1009
DB	301	PhePheProAlaSerSerArgAspGlnSerSerTyrLeuGlnGlyPheSerArgAsnThr	320
QY	1010	TTTGAGGGCCGCTTCAATCGCGAAATTCATAGATACGAGGGTGTGTTAGAGAGAAT	1069
DB	321	LeuGluAlaAlaPheAsnAlaGluPheAsnGluIleArgArgValLeuLeuGluGluAsn	340
QY	1070	GCAGAGGTGACCAAGAGAGAGAGCGGAGCGGATGGAGTACTCGGAGTAGTGAGAAC	1129
DB	341	AlaGlyGlyGluGlnGluArgGlyGlnArgArgTrpSerThrArgSerSerGluAsn	360
QY	1130	RATGAGGAGTATAGTCAAGTGTCAAGAGGACGCTGAGAGACTTACTTAAGCAGCT	1189
DB	361	AsnGluGlyValIleValLysValSerLysGluHisValGluGluLeuThrLysHisAla	380
QY	1190	AAATCCGCTCTCAAGAAAAGGCTCCGAAGAAGAGGAGATATCACCAACCCAACTTGTG	1249
DB	381	LysSerValSerLysLysGlySerGluGluGluGlyAspIleThrAsnProIleAsnLeu	400
QY	1250	AGAGAAGCGGAGCCGATCTTTTAACTTTGGGAAAGTTATTTGAGGTGAAGCCAGAC	1309
DB	401	ArgGluGlyGluProAspLeuSerAsnAsnPheGlyLysLeuPheGluValLysProAsp	420
QY	1310	RAGAGAACCCCGCTTCAGACCTGCAGCATGATGCTCACCTGTGTAGAGATCAAGAA	1369
DB	421	LysLysAsnProGlnLeuGlnAspLeuAspMetMetLeuThrCysValGluIleLysGlu	440
QY	1370	GGAGTTTGATGCTCCACACTTCAACTTCAAGGCGCATGTTATCGTCTCGTCAACAAA	1429
DB	441	GlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleValValAsnLys	460
QY	1430	GGAACGTGAAACCTTGAACTCGTGGCTGTAAAGAAAGCAACAACAGAGGAGCGCGG	1489
DB	461	GlyThrGlyAsnLeuLeuValAlaValArgLysGluGlnGlnArgGlyArgArg	480
QY	1490	GAGAGAGGAGGAG	1549
DB	481	GluGluGluGluAspGluAspGluGluGluGlySerAsnArgGluValArgArgTyr	500
QY	1550	ACAGCGAGTTGAAGAAAGGCGATGTTTCACTCATGCCAGCAGCTCATCCAGTAGCCATC	1609
DB	501	ThrAlaArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIle	520
QY	1610	AACGTTTCTCCGAACTCCATCTGCTTGGCTTCGTTATCAACGCTGAAACCAACACAGA	1669
DB	521	AsnAlaSerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnAsnHisArg	540
QY	1670	ATCTTCTTCGAGGTGATAAGGACCAATGTGATAGACCATAGATAGACCAAGCGAGGAT	1729
DB	541	IlePheLeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAsp	560
QY	1730	TTAGCATTTCCCTGGGTCGGGTGAACAAGTTGAGAAAGCTCATCAAAAACAGAGGAATCT	1789
DB	561	LeuAlaPheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSer	580
QY	1790	CACTTTGTGAGTCTCGTCTCAATCTCAATCTCAATCTCCGCTCTCTCTGAGAAAGAG	1849
DB	581	HisPheValSerAlaArgProGlnSerGlnSerGlnSerProSerSerProGluLysGlu	600
QY	1850	TCTCTGAGAAAGAGGATCAAGAGGAGGAAACCAAGGAGGAGGAGGAGGAGGAGGAG	1909
DB	601	SerProGluLysGluAspGlnGluGlnGlnGlyGlyLysGlyProLeuLeuSer	620
QY	1910	ATTTTGAAGGCTTTTAAAC 1927	
DB	621	IleLeuLysAlaPheAsn 626	

RESULT 3

US-10-508-263-94

[illegible]


```
QY 1685 GATAAGACAAT-----GTGATACACAGATAGAGAACGCG 1723
|||||:|||||
Db 519 AsPLyGluAsnIleSerLySgiuAsnAspValLeuAspGluLySgiuGluAla 538
|||||:|||||
QY 1724 AAGGATTTAGCATTCCTGGGTGGGTGAACAAAGTTGAGAAGCTCATCAAAAAACCAAGAAG 1783
|||||:|||||
Db 539 GluGluThr-----GluGluGluGluLeuGluGluLySAsnGluGlu 552
|||||:|||||
QY 1784 GAATCTCACTTTGTGAGTGTCTCTCAATCTCAATCTCAATCTCAATCTCGTCTCTGAG 1843
|||||:|||||
Db 553 Glu-----ThrGluSerGluIleSerGluAsp 561
|||||:|||||
QY 1844 AAGAGTCTCTGAGAAAGAGGATCAAGAGGAGGAAACCAAGGAGGAAG 1894
|||||:|||||
Db 562 GluGluGluGluGluGluGluGluGluGluGluGluGluGluLySlys 578
|||||:|||||
RESULT 9
US-11-087-099-1159
; Sequence 1159, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 1159
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-11-087-099-1159
Alignment Scores:
Pred. No.: 4,01e-05 Length: 499
Score: 162.50 Matches: 54
Percent Similarity: 45.7% Conservative: 51
Best Local Similarity: 23.5% Mismatches: 81
Query Match: 4.5% Indels: 44
DB: 11 Gaps: 6
US-10-728-323-1 (1-2032) x US-11-087-099-1159 (1-499)
QY 1252 AGAAGCGAGCCGATCTTCTAACACTTTGGAGATTATTGAGTGAGCCAGAC-- 1309
|||||:|||||
Db 57 ArgArgAsnLysLySlysSerSerGlnLeuPheAsnVal-PheGlnGluLySProAsp 76
|||||:|||||
QY 1310 -----AAGAAGAACCCCGAGCTTCAGGA 1332
|||||:|||||
Db 76 eGluAsnCysAsnGlyTrpSerThrValIleAsnArgLysLeuProAlaLeuLySGL 96
|||||:|||||
QY 1333 CTGGACATGATGCTACCTGCTAGAGATCAAGAGAGCTTTGATGCTCCACACTT 1392
|||||:|||||
Db 96 ySerGlnIleGlyIleTyrValValAsnLeuThrLySlysSerMetMetGlyProHisTr 116
|||||:|||||
QY 1393 CAACTCAAGCCATGTTATGCTGCTGCTCAACAAGAACTGGAACCTTGAACCTCGT 1452
|||||:|||||
Db 116 pAsnProMetAlaThrGluIleGlyIleAlaIleGlnGlyGluGlyMetValArgValVa 136
|||||:|||||
QY 1453 GGCTGTAAAGAAAGAGCAACAACAGAGCGGAGCGCGGGAAGAGAGGAGCAAGACGA 1512
|||||:|||||
Db 136 lCys----- 137
|||||:|||||
QY 1513 AGAAGAGGAGGAAGTAAACAGAGAGTGCTAGTACACAGCAGGTGTTGAAGAGCGGA 1572
|||||:|||||
Db 138 -SerLySerGlyThrGlyCysLySAsnMetArgPhe-----LysValGluGluLyAs 155
|||||:|||||
QY 1573 TGTGTTTCATATCCACGACGCTCATCCAGTAGGCATCAACGCT-----TCCTCCGAACT 1626
|||||:|||||
Db 155 pValPheValValProArgPheAspProMetAlaGlnMetAlaPheAsnAsnSerPh 175
|||||:|||||
QY 1627 CCATCTCTTGGCTTTCGTTCAACGCTGAAACACACACAGAAATCTTCTTCGAGGTGA 1686
|||||:|||||
```

```
Db 175 eValPheValGlyPheSerThrThrThrLySlysHisHisProGlnTyrLeuThrGly-- 194
|||||:|||||
QY 1687 TAAGACAAATGTATAGACCATAGATAGAACGAAGGAGATTATTAGCATTCCTCGGTC 1746
|||||:|||||
Db 195 -LysAlaSerValLeuArgThrLeuAspArgGlnIleLeuGluAlaSerPheAsnValGl 214
|||||:|||||
QY 1747 GGGTGAACAAGTTGAGAAGCTCATCAAAAAACCAAGGAATCTCACTTTGTGAGTCTCG 1806
|||||:|||||
Db 214 yAsnThrThrMetHisGlnIleLeuGluAlaGlnGlyAspSerValIleLeu----- 231
|||||:|||||
QY 1807 TCTCAATCTCAATCTCAATCTCCGTCGTCCTCGAGAAAGAGTCTCTCGAGAAAGAGA 1866
|||||:|||||
Db 232 -----GluCysThrSerCysAlaGluGluLySArgLeuMetGluGluGluMe 248
|||||:|||||
QY 1867 TCAAGAGGAGGAAAAACCAAGGAGGGAAG 1894
|||||:|||||
Db 248 tArgLySgiuGluGluGluAlaLySlys 257
|||||:|||||
RESULT 10
US-10-784-004-709
; Sequence 709, Application US/10784004
; Publication No. US20060084066A1
; GENERAL INFORMATION:
; APPLICANT: Biogen Idex
; TITLE OF INVENTION: Surrogate Markers of Pain
; FILE REFERENCE: 08201.6029-00000
; CURRENT APPLICATION NUMBER: US/10/784,004
; CURRENT FILING DATE: 2004-02-20
; NUMBER OF SEQ ID NOS: 1251
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 709
; LENGTH: 5689
; TYPE: PRT
; ORGANISM: human
US-10-784-004-709
Alignment Scores:
Pred. No.: 0.000117 Length: 5689
Score: 162.00 Matches: 153
Percent Similarity: 32.0% Conservative: 87
Best Local Similarity: 20.4% Mismatches: 299
Query Match: 4.5% Indels: 211
DB: 9 Gaps: 28
US-10-728-323-1 (1-2032) x US-10-784-004-709 (1-5689)
QY 17 CATCAATCATCTATATAAGTAGTAGCAGGAGCAATGAGAGGAGGTTTCTCACTGATG 76
|||||:|||||
Db 2727 HisGlnGluSerAlaArgAspTrpSerGlyGlyArgSerGlyArgSerGlySerPheIle 2746
|||||:|||||
QY 77 CTGTTCTAGGGATCCTTCTCTGCTTCAGTTTCTGCAACGCATCCCAAGTCATCACCT 136
|||||:|||||
Db 2747 TyrGlnVal-----SerThrHisGluGlnSerGluSer 2757
|||||:|||||
QY 137 TACCAGAAGAAACACAGAAACCCCTCGCCAGAGTGCTCCAGAGTTGTCAACAGGAA 196
|||||:|||||
Db 2758 AlaHisGlyArgThrArgThrSerThrGlyArgArgGlnGlySerHisHisGluGlnAla 2777
|||||:|||||
QY 197 CCGATGACTTGAAGCAAAAGGCATCGATCTCGTCGCCCAAGCTCCAGTATGATCTCT 256
|||||:|||||
Db 2778 ArgAspSerSerArgHisSerAlaSerGlnGluGlnAspThrIleArgAlaHisPro 2797
|||||:|||||
QY 257 -----CGTTGTCTATGATCTTCGAGGA 280
|||||:|||||
Db 2798 GlySerArgArgGlyGlyArgGlnGlySerHisHisGlnSerValAspArgSerGly 2817
|||||:|||||
QY 281 CACATCTGGC-----ACCACCAACCAAC-- 301
|||||:|||||
Db 2818 HisSerGlySerHisHisSerHisThrSerGlnGlyArgSerHisPalaSerHisGly 2837
|||||:|||||
QY 302 -----CGTTCCCTCCAGGGAGCGGACAGCTGCGCCGCCCAACCCCGAGAC--- 346
|||||:|||||
Db 2838 GlnSerGlySerArgSerAlaSerArgGlnThrArgLySAspLySglnSerGlyAspGly 2857
|||||:|||||
```

Qy	347	-----TACGATGATACCCGCGT	364
Db	2858	SerArgHisSerGlySerArgHisHisGluAlaAlaSerTrpAlaAspSerSerArgHis	2877
Qy	365	CAACCCGAGAGAGAGAGAGAGCCGATGGGACCACTGCACCGA	415
Db	2878	SerGlnValGlyGlnGlu-GlnSerGlySerArgThrSerArgHisGlnGlySerSer	2897
Qy	416	CGTGAAGAG-----AAGAAGACTGGAGACAACCAAGAGAAAGATTGGAGCGCA	463
Db	2897	rValSerGlnAspSerAspSerGluArgHisSerAspAspSerGluArgIleuSerGlySer	2917
Qy	464	CCAAGTCATCAGCAGCCACGAAATA-AGGCCCGAGAGAGAGAGGAGAAACAAGATG	522
Db	2917	rAlaSerArgAsnHisHisGlySerSerArgGluGlnSerArgAspGlySerArgHis--	2936
Qy	523	GGGAACACAGGTAGCATGTGAGGGAAGAAACATCTCGGAACAACCCCTTCTACTTCCC	582
Db	2937	-----ProGlyPheHisGlnGluAspAlaSerHisGlyHis-----	2949
Qy	583	GTCAAGCGGTTTAGCACCCGCTACGGGACCAACCAACGGTAGGATCCGGGTCCTGCAGAG	642
Db	2950	-----SerAlaAspSerSerArgGlnSerGly-----ThrHI	2960
Qy	643	GTTTGACCAAGGTCAGGCGATTTCAGAATCTCCAGAATCACCGTATTGTGCAGATCGA	702
Db	2960	sHisThrGluSerSerHisGlyGlnAlaValSerSerHisGlu-----GlnAlaAr	2978
Qy	703	GGCCAAACCTAACTCTTGTCTTCCCAAGCAC-----GCTGATGCTGATTAACAT	753
Db	2978	gSerSerProGlyGluArgHisGlySerArgHisGlnGlnSerAlaAspSerSerArgHI	2998
Qy	754	CCTGTATTACAGCAAGGCAAGCCCGTCACGTCAGCAATGGCAATTAACAGAAAGAG	813
Db	2998	sSerGlyIleGlyHisArgGlnAlaSerSerAlaValArgAspSerGlyHisArgGlySer	3018
Qy	814	C-----TTTAATCTTGACGAGGCGCATGTCACATCAGAAATCCATCCCGTTTCAT	861
Db	3018	rSerGlySerGlnValThrAsnSerGluGlyHisSer-----	3030
Qy	862	TTCCTACATCTTGAAACCCCATGACACCAACCAACCTTCAGAGTAGCTAAAATCTCCATGCC	921
Db	3031	-----GluAspSerArgThrGlnSerValSerAlaHis-----	3041
Qy	922	CGTTAACACACCCGCGCAGTTTGAGGATTCTTCCCGCGAGCAGCGACCAACATCATC	981
Db	3042	---GlyGlnAlaGlyProHisGlnGlnSerHisLysGluSerAlaArgGlyGlnSerGl	3060
Qy	982	CTACTTCGAGGCTTCAGCGAATACGTTGGAGCGCCCTTCAATCGGAATTCATGA	1041
Db	3060	yGluSerSerGlyArgSerArgSerPheLeuTyrglnValSerSerHisGlnGlnSerGl	3080
Qy	1042	GATACGAGGGTGTGTTAGAAAGAATGCGAGAGGTGACGACAGAGAGAGA---GGCA	1098
Db	3080	uSerThrHisGlyGlnThrAlaProSerThrGlyGlyArgGlnGlySerArgHisGluGl	3100
Qy	1099	GAGCGGATGGAGTACTCGAGTACTGAGACAATGAAGGAGTGTAGTCAAAAGTGTCAA	1158
Db	3100	nAlaArgAsnSerSerArgHisSerAlaSerGlnAspGly-----	3113
Qy	1159	GGACACGCTGAAGAACTTACTAGCACGCTAAATCGCTTCAAAAGAAAGCTCCGAAGA	1218
Db	3114	-----GlnAspThrIleArgGlyHisProGlySerSerArgGlyGly--Arg	3128
Qy	1219	AGGGGAGATATCAACCAACCAATCAACTTGAGAGAAGGGAGCCCATCTTTCTAACAA	1278
Db	3129	GlnGlySerTyHisGlu-----GlnSerValAspArgSerGlyHisSer-----	3143
Qy	1279	CTTTGGGAAGTATTGTGAGGTGAAGCCAGACAGCAAGAAACCCCA-----	1323
Db	3144	-----GlyTyHisHisSerHisThrThrProGlnGlnValArgSerAsp	3157

[illegible]

[illegible]

308 ArgArgAsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisasnIleGlyGln 327

1262 -----CCCGATCTTTCTTAACAACACTTTGGGAAGTTATTAGTGGTGAAGCCAGACAAG 1312

328 ThrSerSerProAspIleTyrAsnProGlnAlaGlySerValThrThrAlaThrSerLeu 347

1313 AAGAACCCCGAGCTTCAGGACCTGGACATGATGCTCACCTGTGTAGAGATCAAGAAGGA 1372

348 AspPheProAlaLeuSerTrpLeuArgLeuSerAlaGluPheGlySerLeuArgLysAsn 367

1373 GCCTTGTATGCTCCACACACTCAACTCAAGAGCCATGGTTATCTGCTGCTGCATCAACAAGGA 1432

368 AlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeuAsnGly 387

1433 ACTGGAAACCTTGAACTCGTGGCTGTAAAGAAAGACCAACACAGAGGGGACGCGGGAA 1492

388 ArgAlaLeuIleGlnValValAsnCysAsnGlyGlu----- 399

1493 GAAGAGGAGGACGAAGACGAAGAGGAGGAAAGTAAACAGAGAGGTGCGTAGGTACACA 1552

400 ----- 403

1553 GCGAGGTGAAGAGGCGATGTGTCATCATGCTCCAGCAGCTCATCTCAGTAGCCATCAAC 1612

404 GlyGluLeuGlnGluGlyArgValLeuIleValProGlnAsnPheValValAlaAlaArg 423

1613 GCCTTCCTCCGAA---CTCCATCTGCTGGCTTCGGTATCAACGCTGAAAAACACCACAGA 1669

424 SerGlnSerAsnPheGluTyrValSerPheLysThrAsn-----AspThrPromet 441

1670 ATCTTCTCTCGAGGTGATAGGACAATGTGATAGACCAAGATAGAGAAGCAAGCAAGGAT 1729

442 IleGlyThrLeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGln 461

1730 TTAGCATTCCTCGGTGGTGAACAAGTTGAGAAGCTCATCAAAAACCAAGAGGAATCT 1789

462 HisThrPheAsnLeuLysSerGlnGlnAlaArgGln---IleLysAsnAsnAsnProPhe 480

1790 CACTTTGTAGTGTCTGCTCTCAATCTCAATCTCAATCT 1828

481 LysPheLeuValProGlnGluSerGlnLysArgAla 493

RESULT 12

US-10-982-545-12

; Sequence 12, Application US/10982545

; Publication No. US20050244890A1

; GENERAL INFORMATION:

; APPLICANT: Davies, Huw Alun

; APPLICANT: McGuire, James

; APPLICANT: Simonsen, Anja Hviid

; APPLICANT: Blennow, Kaj

; APPLICANT: Podust, Vladimir

; TITLE: CIPHERGEN BIOSYSTEMS, INC.

; TITLE OF INVENTION: Biomarkers for Alzheimer's Disease

; FILE REFERENCE: 016866-011550US

; CURRENT APPLICATION NUMBER: US/10/982,545

; CURRENT FILING DATE: 2004-11-06

; PRIOR APPLICATION NUMBER: US 60/518,360

; PRIOR FILING DATE: 2003-11-07

; PRIOR APPLICATION NUMBER: US 60/526,753

; PRIOR FILING DATE: 2003-12-02

; PRIOR APPLICATION NUMBER: US 60/546,423

; PRIOR FILING DATE: 2004-02-19

; PRIOR APPLICATION NUMBER: US 60/547,250

; PRIOR FILING DATE: 2004-02-23

; PRIOR APPLICATION NUMBER: US 60/558,896

; PRIOR FILING DATE: 2004-04-02

; PRIOR APPLICATION NUMBER: US 60/572,617

; PRIOR FILING DATE: 2004-05-18

; PRIOR APPLICATION NUMBER: US 60/586,503

; PRIOR FILING DATE: 2004-07-08

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: PatentIn Ver. 2.1

```

; SEQ ID NO 12
; LENGTH: 677
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Chromogranin B precursor
; NAME/KEY: SIGNAL
; LOCATION: (1)..(20)
; OTHER INFORMATION: signal peptide
; NAME/KEY: PEPTIDE
; LOCATION: (21)..(677)
; OTHER INFORMATION: Chromogranin B (Secretogranin I)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (326)..(385)
; OTHER INFORMATION: biomarker peptide 7258 Da, processed fragment of
; OTHER INFORMATION: Chromogranin B
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (341)
; OTHER INFORMATION: Xaa = sulfotyrosine
US-10-982-545-12

Alignment Scores:
Pred. No.: 0.000274 Length: 677
Score: 153.00 Matches: 143
Percent Similarity: 34.3% Conservative: 101
Best Local Similarity: 20.1% Mismatches: 285
Query Match: 4.3% Indels: 183
DB: 9 Gaps: 34

US-10-728-323-1 (1-2032) x US-10-982-545-12 (1-677)
QY 71 CTGATGCTGTTCTAGGATCTTCTGCTGGCTTCAGTTCTGCA----- 115
Db 6 LeuLeuSerLeuLeuGlyAlaValGlyLeuAlaValAlaValAsnSerMetProValAspAsn 25
QY 116 -----ACGCATGCTCAAGTCATCACCTTAC 139
Db 26 ArgAsnHisAsnGluGlyMetValThrArgCysIleIleGluValLeuSerAsnAlaLeu 45
QY 140 CAGAGAAACAGAGAACCCCTGCGCCACAGAGTGCTCCAGAGTTGTCAACAGGAACCG 199
Db 46 SerLysSerSerAlaProIleThrProGluCysArgGlnValLeuLysThrSerArg 65
QY 200 GATGACTTGAACAAAG---GCATGCGAGTCTCGTCGCACCAAGCTCGATGATCCT 256
Db 66 LysAspValLysAspLysGluThrThrGluAsnGluAsnThrLysPheGluVal----- 83
QY 257 CTTTGTGCTATGATCTCGAGGACACACTGCGCACCAACCAACGTTCCCTCCAGG 316
Db 84 ArgLeuLeuArgAspPro---AlaAspAlaSerGluAlaHisGluSerSerSerArgGly 102
QY 317 GAGCGGACACGTGGCGCCCAACCCGAGACTACGATGATGACCGCGCTCAACCCCGAAGA 376
Db 103 Glu-----AlaGlyAlaProGly-----GluGluAspIleGlnGlyProThrLys 117
QY 377 GAGGAAGGCGCGATGGGGACACAGTGGACCGAGGGAGCGGTGAAGA-----GAAGAA 430
Db 118 AlaAspThrGluLysTrpAlaGluGlyGlyGlyHisSerArgGluArgAlaAspGluPro 137
QY 431 GACTCG-----AGACAACCAAGAGAGATTGGAGCGGACCAAGTCAT 472
Db 138 GlnTrpSerLeuTyProSerAspSerGlnValSerGluGluValLysThrArgHisSer 157
QY 473 CAGCAGCCACCGAAATAGGCCCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 520
Db 158 GluLysSerGlnArgGluAspGluGluGluGluGluGluGluGluGluGluGluGluGlu 177
QY 521 TGGGGAACACCGAGGTAGC-----CATGTAGGGGAAGAAACATCTCGGAAACACCT 571

```

```

Db 178 ArgGlyGluAspSerSerGluGluLysHisLeuGluGluProGlyGluThrGlnAsnAla 197
QY 572 TTCTACTTCCCGTCAAGCGGGTTTAGCACC----- 601
Db 198 PheLeuAsnGluArgLysGlnAlaSerAlaIleLysLysGluLeuValaAlaArgSer 217
QY 602 -----CGCTACGGAAACAAACCGTAGGATCCGGGTCTCGACAGAGTTTACCAA 652
Db 218 GluThrHisAlaAlaGlyHisSerGlnGlyThrHisSerArgGluLysSerSerGln 237
QY 653 AGGTCAAGCGCAGTTTCAGAATCTCAGAATCACCGTATTGTGCAGATCGAGGCCAACCT 712
Db 238 GluSerGlyGluGluAlaGlySerGlnGluAsnHisProGlnGluSerLysGlyGlnPro 257
QY 713 AACACTTGTCTTCTCCCAAGCACGCTGATGCTGATAACATCTCTTGTATTATCCAGCAAGG 772
Db 258 ArgSerGlnGluGluSerGluGluGlyGluAlaSerAlaThrSerGluValAspLysArg 277
QY 773 CAAGCCACCGTACCGTAGCAATGGCAATTAACAGAAAGAGCTTTAACTTTGACGAGGCG 832
Db 278 ArgThrArgProArgHisHisGlyArgSerArgProAspArgSerSerGlnGlyGly 297
QY 833 CATGCACCTCAGATCCCATCCGTTTCATTCTTCTACATCTTGAACCGCCATCACACACAG 892
Db 298 -----SerLeuProSerGluGluLysGlyHisProGlnGluGluSerGluGluSer 314
QY 893 AACCTCAGATAGCTAAATCTCCATGCCGTTAAACACACCCGCGCAGTTTGAAGATTTC 952
Db 315 AsnValSerMetAlaSerLeu-----GlyGluLysArgAspHis 327
QY 953 TTCCCGCGCAGCAGCCGA-----GACCAATCATCTTACTTTCGAGGCG 994
Db 328 HisSerThrHisTyArgAlaSerGluGluGluProGlu**GlyGluGluLysGly 347
QY 995 TTCAGCAGATACGTTGAGCGCCCTTCAATCGGAATTCATAGATAGATAGCGAGGTTG 1054
Db 348 TyrPro-----GlyValGlnAlaProGluAspLeuGluLysTrpGluArgTyrArg----- 363
QY 1055 CTGTTAGAAAGAAATGCGAGGAGTGAGCAAGAGAGAGAGAGGCGCAGAG----- 1102
Db 364 -----GlyArgGlySerGluGluTyArgAlaProArgProGlnSerGlu 378
QY 1103 ---CGATGAGTACTCGGAGTAGTGAGAACAAATGAAGAGGTAGTATAGTCAAGTGTCAAG 1159
Db 379 GluSerTrpAspGluGluAspLysArgAsnTyProSerLeu----- 392
QY 1160 GAGCAGCTTGAAGAACTTACTAAGCACCGCTAAATCCGCTCAAGAAAGGCTCCGAAGA 1219
Db 393 -----GluLeuAspLysMetAlaHisGlyTyArgGlyGluGlu---SerGluGlu 407
QY 1220 GAGGAGATATACCAACCCCAATCACTTGAGAGAAAGCGCAGCCGATCTTTCTAACAC 1279
Db 408 Glu--ArgGlyLeuGluProGlyLysGlyArgHisHisArgGlyArgGly----- 423
QY 1280 TTTGGGAAGTTATTGAGGTGAAGCCAGA-----CAAGAAGAAC 1318
Db 424 -----GlyGluProArgAlaTyProMetSerAspThrArgGluGlu 438
QY 1319 CCCAGCTTTCAGGACCTGACATGCTCACCTGTGTAGATATCAAGAGAGGAGCTTTG 1378
Db 438 ysaArgPheLeuGlyGlyHisHis-----ArgValGlnGluAsnGlnMetA 454
QY 1379 ATGCTCCACACTTCAACTCAAGGCCATGGTTATCGTCGTCGTCACAAAGAGAACTGA 1438
Db 454 spLysAla-----ArgArgHisProGlnGlyAlaTrpL 465
QY 1439 AACCTTGAACCTCGT-----GGCTGTAAAGAAAGAGCAACAA 1474
Db 465 ysaGluLeuAspArgAsnTyProGlyGlyGluGlyAlaProGlyLys--TrpGln 484
QY 1475 CAGAGGCGACCGCGGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1534
Db 485 GlnGlnGlyAspLeuGlnAspThrLysGluAsnArgGluGluAlaArgPheGlnAspLys 504

```

QY 1535 GAGGTG---CGTAGGTACACAGG-----AGTTGAAGAAGGCGCATGTGTTTCATCATG 1585
Db 505 GlnTyrSerSerHisHisThrAlaGluLysArgLysArgLysGluLeuPhe-----Asn 523
QY 1586 CCAGCAGCTCATCCAGTAGCATCAACAGCTTCTCCGAACTCCATCTGCTTGGCTTCGGT 1645
Db 524 ProTyrTyrAspProLeuGlnTrpLysSerSer-----HisPhe----- 536
QY 1646 ATCAACGCTGAAACCAACCAAGAAATCTTCTTCAGAGGTGATAAGCAACATGTATAGAC 1705
Db 537 GluArgArgAspAsnMetAsnAspAsnPheLeuGluGlyGluGluGluAsnGluLeuThr 556
QY 1706 CAGATAGAGAAG----- 1717
Db 557 LeuAsnGluLysAsnPhePheProGluTyrAsnTyrAspTrpTrpGluLysLysProPhe 576
QY 1718 -----CAAGCGAAGGATTTAGCATTCCTCGGTGGGTGAA 1753
Db 577 SerGluAspValAsnTrpGlyTyrGluLysArgAsnLeuAla----- 590
QY 1754 CRAAGTTGAGAAGTCT---ATCAAAAAACCAAGAGAA-----TCTACTTTT 1795
Db 591 ArgValProLysLeuAspLeuLysArgGlnTyrAspArgValAlaGlnLeuAspGlnLeu 610
QY 1796 GTGAGTGTGCTCTCAATCTCAATCTCAATCTCGTCTGCTCTCTGAGAAAGAGTCTCT 1855
Db 611 LeuHisTyrArgLysLysSer---AlaGluPheProAspPheTyrAspSerGluGluPro 629
QY 1856 GAGAAAGAGGATCAAGCAGGAGGAGAAACCAA 1885
Db 630 ValSerThrHisGlnGluAlaGluAsnGlu 639

RESULT 13

US-11-072-512-3378
; Sequence 3378, Application US/11072512
; Publication No. US20060029945A1

GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3378
; LENGTH: 623

; TYPE: PRT

; ORGANISM: Homo sapiens

US-11-072-512-3378

Alignment Scores:

Pred. No.: 0.000566 Length: 623
Score: 149.00 Matches: 134
Percent Similarity: 34.2% Conservatives: 97
Best Local Similarity: 19.9% Mismatches: 274
Query Match: 4.2% Indels: 170
DB: 11 Gaps: 34

US-10-728-323-1 (1-2032) x US-11-072-512-3378 (1-623)

QY 116 ACGCATGCCAAGTCATCATCCCTTACCAGAAAGAAAACAGAACCCCTGCCCCAGAGGTGC 175
Db 5 SerHisPheSerGlnProProTyrArgAspLeuTrpProProArgProGlyGlyGlu 24
QY 176 CTCACAGAGTTGCAACAG-----GAACCGGAGTACTTGAAGCAAAAGGCATCC--- 223
Db 25 ArgGluSerThrGlnArgLeuGlyGlnArgSerGlyAlaAspSerThrAlaCysSer 44
QY 224 -----GAGTCTCGCTCCACCAAGCTCGAGTATGATCTCTCTTGT 262
Db 45 ArgAlaGlyThrProGlyAlaGluSerGluAlaGlyAlaCysTrpLeuHisProHisCys 64
QY 263 GTCTATGATCTCT-----CGAGGACACACTGGC 289
Db 65 SerPheThrProArgProArgArgGlyCysSerAspSerLeuArgGlySerArgSer 84
QY 290 ACCACCAACCAAGTTCCTCCAGGGAGCGGACGACAGTGGCGCCCAACCCGGAGACTAC 349
Db 85 LeuSerAspValAlaArgArgProLeuGluArgSerArgLysHisArgPro----- 101
QY 350 GATGATGACCGCGTCAACCCCAAGAGAGAGGAGGCGCGATGGGACAGCTGGACCG 409
Db 102 -----ArgSerArgArgLeuGluAspAla-----TrpGlyGluThrGlyThr 115
QY 410 AGGAGCGGTGAAGAGAAAGACTGGAGACAAACCAAGAGAGATGGAGGCGACCAAGT 469
Db 116 Lys-----ProArgProAlaTrpGlnProGlnThr 125
QY 470 CATCAGCAGCCACGGAAATAAGGCC----- 496
Db 126 GlnLeuProGlnArgProGlnProCysProHisTyrProLeuAlaGlnGlyAspSer 145
QY 497 -----GAAGGA 502
Db 146 ProProProCysProGlyGlyAlaGlyThrProLeuSerGlyThrPheArgValGluLys 165
QY 503 AGAAGAGAGAACAGAGTGGGGAACACCA---GGTAGCCATCTGAGGAGAAACATCT 559
Db 166 AlaGlnGlyGlyAspGlnTrpAlaValProLeuGlyArgHisLeu-----Gly 181
QY 560 CGGAACAACCTTTCTTCTTCTCCGTCACAGCGGCTTTAGCACCC-----CGTACGGG 610
Db 182 ArgTrpSerProSerSerValProSerGluArgSerSerValProSerGlnLysPheLys 201
QY 611 AACCAAAACGGTAGGATCCGGGTCTCTGCAGAG-----TTTGACCAA----- 652
Db 202 ArgHisSerAlaCysValCysAlaGlnLysArgAspSerSerAspGlnValGluSerLeu 221
QY 653 AGGTCAAGGAGTTTCAGAAATCTCCAGAAATCACCCTGATTGTGCAGATCGAGGCCAACCT 712
Db 222 AlaSerArgAspSerGlnProLeuAlaSerSerLysGluMetArg-----SerPro 238
QY 713 AACACTCTTGTCTCTCCCAAGCAGCGTGTGCTGATAACATCTCTGTTATCTCCAGCAGGG 772
Db 239 HisThrGlnValLeu-----LysSerLysLeuGluGluValValSerSerGlnAsp 256
QY 773 CAAGCCACCGTGCACCGTAGCAATGGCAATAACAGAAAGAGCTTTAATCTTGTACGAGGGC 832
Db 257 GlnGlnIleValAlaLeuValLeuThrArgLeuLysLysAlaGlnArgIleArgGlu--- 275
QY 833 CATGCACTCAGAATCCATCCCGGTTTCAATTCCTACATCTTGAACCGCCATGACACACAG 892
Db 276 -----LeuGlnGlnGlnAlaAlaLysAlaTrpGluGluLysArgSerAspGln--- 292

```
QY 893 AACCTCAGAGTAGTAAATCTCCATGCCCCGTTAAACACACCCCGCCAGCTTTGAGGATTC 952
Db 293 -----LysValGlnMetThrLeuGluArgLeu----- 304
QY 953 TTCCCGCGCAGCAGCAGACCAATCATCTCTACTTCAGGGCTTCAGAGGATACGTTG 1012
Db 305 -----LeuLeuArgGlnSerGlnGluGlnTrpGlnGluGluGlnArgLysThrLeu 322
QY 1013 GAGGCGCGCTCAATCGCAATCAATGAGTACGAGGGTGCTGTGTAGAACAGATGCA 1072
Db 323 Gln-----SerProGluGlnArgGlyLeuArgAspSerGlnArgLysAsnVal 339
QY 1073 GGAGGTGACAAAGAGAGAGAGCGCAGAGCGGATGAGTACTCGGAGTAGTGAGAACAA 1132
Db 340 -----ProProGlyGluSerArgTrpLysGluGlnProGluAspGlnGlu 354
QY 1133 GAAGGAGTAGTAGTCAAGTGTCAAGGAGCAGCGTTGAAGAACTTACTAAGCAGCGTAA 1192
Db 355 SerProArgGlnGluLysLeuGluLysAlaArgAlaGln-----AlaGluHisArgLys 372
QY 1193 TCCGCTCAAAAGAAAGCGCTCCGAAGAGAGAGGAGATATCACCAACCACTCAACTTGAG 1252
Db 373 GlnCysGlnValArgLeuArgGluGlnGluGlnLysMetLeuArg-----AsnLeuArg 390
QY 1253 GAAGCGAGCCC---GATCTTTCTAAACAACTTTGGGAAG-----TTA 1291
Db 391 GlnGlnHisSerLeuGlnLeuGlnArgArgLeuValGluAlaCysArgLysArgHisLeu 410
QY 1292 TTTGAGGTGAGCCAGACAAAGAAACCCCGCTTCAGGACTTGACATG----- 1342
Db 411 HisAlaValGluGlyGlnLysLys-----ValGlnAspThrAsnLeuSerSerLeu 427
QY 1343 -----ATGCTCACCTGTGTAGATCAAAAGAGAGGCTTTG 1378
Db 428 IleAsnTyxGlnAlaArgLysValLeuMetAspCysGlnAlaLysAlaGluGluLeuLeu 447
QY 1379 ATGCTCCACACTTCAACTCAAGGCCATCGTTATCGTCTCGTCAACAAAGAACTGGA 1438
Db 448 ArgGlnLeuSerLeuGluGlnSerPheGlnArgSerGlnGluLeu----- 465
QY 1439 AACCTTGAAGTCTGCTGTAGAAAGACCAACACAGGGGCGGCGGAAGAGAG 1498
Db 466 -----LeuArgLysGluArgGlnArgLeuArgGluLysAlaGln 479
QY 1499 GAGGACGACGACGAGAGAG-----GGAAGTAACAGAGGTGCT 1543
Db 480 LysGluGluGlnGlnGlnAlaArgTrpArgAlaGlyGluSerGluGluGlnArg 499
QY 1544 AGGTACACAGCGAGGTG-----AAGGAAGCGATGTGTTTCATATGCCAGCA---GCT 1594
Db 500 LysMetArgLysArgIleLeuValGluLeuAlaAspGluLysIleArgGlnAlaArgSer 519
QY 1595 CATCCAGTAGCCATCAACCTCTCCGAACTCCATCTCTGCTGGCTCGGTATCAACGCT 1654
Db 520 HisValHisLysThrThrArgAspLysValGlnHisLeuArgGluLeuAsnHisLeuArg 539
QY 1655 GAAACACACCAAGAGATTTAGCATTTCCCTGGGTGGGTGAACAAGTTGAGAGCTCATCAA 1714
Db 540 GluLysAsnHisIleLeuLysLeu-----LysAlaGlu 551
QY 1715 AAGCAAGCAAGAGATTTAGCATTTCCCTGGGTGGGTGAACAAGTTGAGAGCTCATCAA 1774
Db 552 LysGluGluLys---CysHisIleGluGlyIleLysGluAlaIleLysLys-----Lys 568
QY 1775 AACCAGAAGAACTCTACATTTGTGAGTGTCTGCTCAATCTCAATCTCAATCTCGGTG 1834
Db 569 GluGlnArgValGlnHisIle-----SerGlnGlyLysAspProAsn 582
QY 1835 TCTCTCTGAG---AAAGAGTCTCTGAGAAAGAGGATCAAGAGAG 1876
Db 583 PheGlnGluPheGlnLysLeuProGlnAlaSerArgArgGluGlu 597
```

RESULT 14

```
US-10-508-263-24
; Sequence 24, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Glycine max
US-10-508-263-24
```

```
Alignment Scores:
Pred. No.: 0.000656 Length: 562
Score: 148.00 Matches: 68
Percent Similarity: 38.8% Conservative: 45
Best Local Similarity: 23.4% Mismatches: 129
Query Match: 4.1% Indels: 49
DB: Gaps: 11
```

US-10-728-323-1 (1-2032) x US-10-508-263-24 (1-562)

```
QY 292 CACCAACCAACGTTCCCTCCAGGGAGCGGACACGTGGCCCAACCCGGAGACTACCA 351
Db 298 HisProProArgArgProSerHisGly-LysArgGluGlnAspGluAspGluAspGlu 317
QY 352 TGATCACCGCCCTCAACCCCGAAGAGAGGAGGAGCGGATGGGACCACAGCTGGACCGAG 411
Db 317 pGluAspLysProArgProSerArgProSerGlnGlyLysArgAsnLysThrGlyGlnAs 337
QY 412 GGAGCGTGAAGAGAAAGAA-----GACTGGAGACAACCAAG 447
Db 337 pGluAspGluAspGluAspGluAspGlnProArgLysSerArgGluTrpArgSerLysLy 357
QY 448 AGAAGATTGGAGCGCAGCAAGTCAACAGCAGCCAGGAAATAAGCCCGAAGAGGAGAGA 507
Db 357 sThrGlnProArgArgProArgGlnGluGluProArgGluArgGlyCysGluThrArgAs 377
QY 508 AGGAGAACAAAGAGTGGGAAACACACAGGTAGCCATGTGGGGAAGAAACATCTCGGAACA 567
Db 377 nGlyValGluGlu---AsnIleCysThrLeuLysLeuHisGluAsnIleAlaArg----- 394
QY 568 CCCTTTTACTTCCGTCAGCGGTTTAGCACCCGCTACGGGAACCAACAAACGGTAGGAT 627
Db 395 -----ProSerArg-----AlaAspPheTyxAsnProLysAlaGlyArgIle 408
QY 628 CCGGTCTCTGACAGAGTTTGACCAAGGTCA---AGGCAGTTTTCAGAAATCTCCAGAATCA 684
Db 408 sSerThrLeuAsnSerLeuThrLeuProAlaLeuArgGlnPheGlnLeuSerAlaGlnTy 428
QY 685 CCGTATTGTGACAGTCGAGGCGCAACCTAAACACTCTTGTCTTCCCAAG---CACGCTGA 741
Db 428 rValValLeuTyx-----LysAsnGlyIleTyxSerProHisTrpAsnLeuAs 444
QY 742 TGCTGATTAACATCTTGTATTATCCAGAGGGCAAGCCCGCTAGCCGTAGCAAAAT----- 796
Db 444 nAlaAsnSerValIleTyxValThrArgGlyGlnGlyLysValArgValValAsnCy8G1 464
QY 797 -GGCAATAACAGAGAGCTTTAATCTTGACGAGGCGCATGCATCAGATCCCATCCCGG 855
Db 464 nGlyAsnAlaValPheAspGlyGluLeuArgArgGlyGlnLeuLeuValValProGlnAs 484
QY 856 TTTTCATTCTTACATCTTGAACCGCCATGACAAACCAACCTCAGAGTAGCTAGCTAAAATCTC 915
Db 484 nPheVal-----ValAl 488
QY 916 CATGCCCGTTAACACACCGCGCAGTTTGTAGGATTTCTTCCGGCGAGCAGCCGAGACCA 975
```



```
QY 1177 -----TACTAAGCAGCTAAATCCGTCCTCAAAAGAAA 1207
Db 434 laArgAspCysSerArgSerArgSerProTyrLysAlaArg---AspArgSerArgSerA 453
QY 1208 GGCTCCGAAGAAGAGGAGATATCAC-----CAACCAATCAACTT---GAGA 1252
Db 453 rgSerProAsnLysAlaArgAspHisSerArgSerArgSerProAsnLysAlaArgAspA 473
QY 1253 GAAGGCGAGCCGATCTTTCTAACAACCTTTGGGAAGTTATTGAGGTGAAGCCAGACAAG 1312
Db 473 rgSerArgSerArgSer-----ProSerLysGluArgA 484
QY 1313 AGAACCCTCCAGCTTCAGACCTTGACATGATCCTCCTGTGTGATGATCAAGAAGGA 1372
Db 484 sphHisSerGlnLeuGlySerProSerLysGlu-----ArgAspHisArgArgS 500
QY 1373 GCTTTGATGCTCCACACTTCAA-----CTCAAAGGCC 1405
Db 500 er---ArgSerProSerLysGluArgGlnCysArgGlnSerArgSerSerLysGluA 519
QY 1406 ATGTTTATCGT-----CGTCGTCAACAAGGAACT 1435
Db 519 rgAspHisArgArgSerArgSerProSerLysGluArgGlnArgGlnSerArg-Ser 538
QY 1436 GGAACCTTTCAACTCGTGGCTGTAGAAAAGAGACAAACAGAGGGGACGGCGGAAGAA 1495
Db 539 ProAsnLysGlu-----ArgAspArgSerGlnSerArgSerProSerGluGlu 554
QY 1496 GAGGAGCAGACAGACAGAGAGGAGGAAGTAAACAGAGAGTGGTAGGTACACAGCG 1555
Db 555 ArgGluHisArgGlnSerArgSerProSerLysGluArgAspArgArgTrpArgSer 574
QY 1556 AGGTTCAAGGAGGCGATGTGTTTCATCATCCACAGCAGCTCATCCAGTAGCCATCAACGCT 1615
Db 575 ProSerLysGluArgGlu-----ArgArgGlnSerArgSer 586
QY 1616 TCCTCCGAACCTCCATCTGCTTGGCTTCGGTATCAACGCTGAAAACCAACACAGAACTTTC 1675
Db 587 SerSerGlu-----GluArgAspHisSerArgSer 596
QY 1676 CTTGCAGGTGATAGACACAATGTGATAGACCAGATAGAGAAGCAACGAGGATTTAGCA 1735
Db 597 ArgSerProAsnLysGlnSerGlyTyrSerArgProArgAlaSerSerLysGluLysAla 616
QY 1736 TTCCCTGGGTGCGGTGAACAAGTTGAGAAGCTCATCAAAAACAGAGGAATCTCACTTT 1795
Db 617 HisSerArgSerArgThrProSerLysGluGlyAsnHisSerGlnSerArgThrSerSer 636
QY 1796 GTGAGTGTCTGCTCCTCAATCTCAATCTCCGTCGTCCTCTCTGAG---AAAGAGTCT 1852
Db 637 LysGluSerAspPro---SerGlnSerThrValProArgSerProAspTrpLysArgSer 655
QY 1853 CTTGAGAAAGAGGATCAAGAGGAGGAACCAA 1885
Db 656 ProThrArgThrSerSerLeuSerGlnAsnArg 666
```

Search completed: May 16, 2006, 00:04:26
Job time : 92.9525 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 21:46:12 ; Search time 13.6006 Seconds

(without alignments)

4312.597 Million cell updates/sec

Title: US-10-728-323-1

Perfect score: 3586

Sequence: 1 aataacatataattcatc.....cgttgtggtgtttcttccc 2032

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapcp 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p model -DEV=xlp
-Q=/abs/ABSSWEB_spool/US10728323/runat_15052006_172134_22413/app_query.fasta_1
-DB=PIR -OFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=us10728323 @CN 1.1 92 @runat_15052006_172134_22413 -NCPU=6 -ICPU=3
-NO MAP -NEG_SCORES=0 -WAIT -DSBBLCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPO=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR 80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1282	35.8	605	2	S20007
2	1281	35.7	605	1	FWSYBA
3	1267	35.3	571	2	S00566
4	1251	34.9	639	2	B24810
5	1178.5	32.9	459	2	S08505
6	1174.5	32.8	433	2	S00567
7	1156	32.2	439	1	FWSYCB
8	1148	32.0	438	2	S35757
9	1109.5	30.9	463	2	A27288
10	1108.5	30.9	463	2	S06309
11	1068	29.8	410	1	FWPMVB
12	962.5	26.8	445	2	JQ2264
13	959.5	26.8	445	2	S00281
14	907.5	25.3	810	2	T44430

15	839	23.4	605	2	S06398	alpha-globulin typ
16	825.5	23.0	436	2	A23498	phaseolin type alp
17	809	22.6	566	2	S22477	vicilin precursor
18	806.5	22.5	430	2	S10156	alpha-phaseolin pr
19	788	22.0	421	2	A24810	phaseolin beta cha
20	786	21.9	509	2	S08059	alpha-globulin typ
21	778	21.7	588	1	FWCNAB	alpha-globulin B p
22	728.5	20.3	275	1	FWPMVA	vicilin A precursor
23	629.5	17.6	582	2	B53234	vicilin-like stora
24	629	17.5	573	2	A53234	globulin-1S, GLB1S
25	614	17.1	537	2	S35221	globulin Begl prec
26	587	16.4	218	2	S16334	beta-conglycinin a
27	569	15.9	450	2	S15675	globulin-2 precurs
28	569	15.9	524	2	JQ1730	62K sucrose-bindin
29	562	15.7	483	2	T06459	62K sucrose-bindin
30	559.5	15.6	448	1	S24756	vicilin-like stora
31	536.5	15.0	240	2	T06399	7S storage protein
32	470.5	13.1	196	2	S02281	convicillin (clone
33	463.5	12.9	540	2	S21825	vicilin-like stora
34	403	11.2	414	2	S60190	vicilin - Zamia fu
35	397.5	11.1	296	2	T06572	convicillin precurs
36	336	9.4	184	1	FSPB	phaseolin - kidney
37	291	8.1	499	2	T10443	probable major pro
38	284	7.9	124	1	FWPMVA	vicilin, 14K compo
39	284	7.9	407	2	T02258	globulin1 - maize
40	269.5	7.5	236	2	T01662	globulin-1 - maize
41	269.5	7.5	699	2	E84565	hypothetical prote
42	263.5	7.3	511	2	E84685	probable seed stor
43	260.5	7.3	481	2	T14300	hypothetical prote
44	228.5	6.4	499	2	D85433	globulin-like prot
45	211.5	5.9	516	1	FWCNBB	beta-globulin B pr

ALIGNMENTS

RESULT 1

S20007

beta-conglycinin alpha chain precursor - soybean

C:Species: Glycine max (soybean)

C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004

C:Accession: S20007

Ribelleve, J.M.; Dickinson, C.D.; Dickinson, L.A.; Nielsen, N.C.

Plant Mol. Biol. 18, 259-274, 1992

A:Title: Synthesis and assembly of soybean beta-conglycinin in vitro.

A:Reference number: S20007; MUID:92119248; PMID:1731988

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-605 <LEL>

A:Cross-references: UNIPROT:Q94LX2; UNIPARC:UPI00000A0D19

C:Superfamily: glycinin

Alignment Scores:

Pred. No.: 1.4e-81 Length: 605
Score: 1282.00 Matches: 285
Percent Similarity: 59.4% Conservative: 108
Best Local Similarity: 43.1% Mismatches: 173
Query Match: 35.8% Indels: 96
DB: 2 Gaps: 18

US-10-728-323-1 (1-2032) x S20007 (1-605)

QY	50	ATGAGAGGAGGGGTTTCTCCACTGATCTGTGTCTAGGATCTCTCTCTGCTTCAGTT	109
DB	2	MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyValPheLeuAlaSerVal	19
QY	110	TTGTCAACGCATGCCAAGTCACCTTACAGAGAAACAGAACCCCTGCGCCAG	169
DB	20	SerValSerPheGlyIleAla-----TyrTrpGluLys---GluAsnProLysHisAsn	36
QY	170	AGTGCTCTCAGAGTGTCAACGAGACCGGATGACTTGAACGAAAGCATGCGAGTCT	229
DB	37	LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla	56

A; Cross-references: UNIPARC:UPI00001745FC; UNIPARC:UPI00001745FD; UNIPARC:UPI00001745FE
 A; Experimental source: seed
 C; Superfamily: glycoprotein
 C; Keywords: glycoprotein; seed; storage protein
 F; 1-22/Domain: signal sequence #status predicted <SIG>
 F; 23-62/Domain: propeptide #status predicted <PRO>
 F; 63-605/Product: beta-conglycinin alpha chain #status predicted <MAT>
 F; 261,517/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
 Pred. No.: 1,65e-81 Length: 605
 Score: 1281.00 Matches: 285
 Percent Similarity: 59.2% Conservative: 107
 Best Local Similarity: 43.1% Mismatches: 174
 Query Match: 35.7% Indels: 96
 DB: 1 Gaps: 18

US-10-728-323-1 (1-2032) x FMSYBA (1-605)

QY 50 ATGAGAGGAGGAGGTTCTCCACTGATGCTGTGCTAGGATCCTGTCTCGCTTCAGTT 109
 DB 2 MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyLeuValPheLeuAlaSerVal 19
 QY 110 TTGCAACGCATGCCAAGTATCACCTTACAGAGAAACAGAGAACCCCTGCGCCAG 169
 DB 20 SerValSerPheGlyIleAla-----TyrTrpGluLys---GluAsnProLysHisAsn 36
 QY 170 AGGTGCTCCAGAGTTGTCAACAGGACCGGATGCTTGAAGCAAAAGGCATGCCAGTCT 229
 DB 37 LysCysLeuGlnSerCysAsnSerGluArgAspSerTyrArgAsnGlnAlaCysHisAla 56
 QY 230 CGCTGCACCAAGCTCGAGTATGAT-----CCTCGTTGT 262
 DB 57 ArgCysAsnLeuLeuLysValGluLysGluLysGluLysGluLysGluLysGluLys 75
 QY 263 GTCTATGATCTCGAGGACACACTGGCACCAACCAACGTTCCCTCCAGGGAGCGG 322
 DB 76 -----ProArgProArg-----ProGlnHisProGlu 84
 QY 323 ACACGTGGCGGCACCCGAGACTACGATGATGAC----- 358
 DB 85 ArgGluProGlnProGlnProGlnProGlnProGlnProGlnProGlnProGlnProGln 104
 QY 359 -----CGCGCTCAACCCCGAAGAGAG-----GAAGGAGCCGA 391
 DB 105 ProPheProArgProGlnProArgGlnGluGluHisGluLysGluLysGluLysGlu 124
 QY 392 TGG-----GGACCGCTGGACCGAGGAGCGTGAAGAGAGAA 430
 DB 125 TrpProArgLysGluLysGluLysGluLysGluLysGluLysGluLysGluLysGlu 144
 QY 431 GACTGGAGACACCAAGAGAGAT-----TGGAGCGGACCAAGTCATCAGCAGCCACGG 484
 DB 145 AspGluGluGlnAspGluArgGlnPheProPheProArgProProHisGlnLysGlu 164
 QY 485 AAAATAAGGCCGCAAG 544
 DB 165 ArgAsnGluGluAspGluAspGluGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 184
 QY 545 AGGGA-----GAAACATCTCGGAACAACCTTTCTACTTCCCGTCAAGCGGCTTAGC 598
 DB 185 SerGluLeuArgArgHisLysAsnLysAsnProPheLeuPheGlySerAsnArgPheGlu 204
 QY 599 ACCCGTACGGNACCAACCGTAGGATCCGGTCTCGAGGTTTACACCAAGAGTCA 658
 DB 205 ThrLeuPheLysAsnGlnTyrGlyArgIleArgValLeuGlnArgPheAsnGlnArgSer 224
 QY 659 AGGCAGTTTCAGAACTCCAGAGATACCGTATTGTGCAGATCGAGGCCAAACCTAACACT 718
 DB 225 ProGlnLeuGlnAsnLeuArgAspTyrArgIleLeuGluPheAsnSerLysProAsnThr 244
 QY 719 CTTGTTCTCCCAAGCACGCTGATGCTGATACATCTCTTGTATCCAGCAAGGCGAAGCC 778

DB 245 LeuLeuLeuProAsnHisAlaAspAlaAspTyrLeuIleValIleLeuAsnGlyThrAla 264
 QY 779 ACCGTGACCGTAGCAAAATGGCAATTAACAGAGAGCTTTAATCTTGACGAGGCCATGCA 838
 DB 265 IleLeuSerLeuValAsnAsnAspArgAspSerTyrArgLeuGlnSerGlyAspAla 284
 QY 839 CTCGAATCCCATCCGGTTTCATTTCTTACATCTTGAACCGCCATGACACCAAGAACCTC 898
 DB 285 LeuArgValProSerGlyThrThrTyrTyrValValAsnProAspAsnAsnGluAsnLeu 304
 QY 899 AGAGTAGCTAAATCTCCATGCCCTTAACACACCCGGCCAGTTTGAGGATTTCTTCCCG 958
 DB 305 ArgLeuIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSerPheLeu 324
 QY 959 GCGAGCAGCCGAGACCAATCATCTTCTGAGGCTTTCAGCAGCAATACGTTGAGGCGC 1018
 DB 325 SerSerThrGluAlaGlnSerTyrLeuGlnGlyPheSerArgAsnIleLeuGluAla 344
 QY 1019 GCCTTCAATCGGAATTCATAGATACGGAGGCTGCTTTAGAGAGAAATCCAGAGGT 1078
 DB 345 SerTyrAspThrLysPheGluGluIleAsnLysValLeuPheSerArgGluGly--- 363
 QY 1079 GAGCAAGAGGAGAGGCGAGCGATGCGAGTACTCGAGTAGTGAGAGACAATGAAGA 1138
 DB 364 -----GlnGlnGlnGlyGluArgLeu-----GlnGluSer 374
 QY 1139 GTGATAGTCAAGTGTCAAAGGAGCAGCTTGAAGAACTTACTTAAGCAGCTAAATCCGTC 1198
 DB 375 ValIleValGluIleSerLysGluGlnIleArgAlaLeuSerLysArgAlaLysSerSer 394
 QY 1199 TCAAGAAAGGCTCCGAAGAGAGGAGATATACCAACCCCAATCAACTTGAGAGAGGC 1258
 DB 395 SerArgLysThrIleSerSerGlu-----AspLysProPheAsnLeuArgSerArg 411
 QY 1259 GAGCCGATCTTTCTAAACACTTTGCGAAGTATTATTGAGGTGAACCCAGAGAGAGAAC 1318
 DB 412 AspProIleTyrSerAsnLysLeuGlyLysPheGluIleThrProGlu---LysAsn 430
 QY 1319 CCCAGCTTCAGGACCTGACATGATGCTCACTGCTGTAGAGATCAAGAGAGAGCTTGT 1378
 DB 431 ProGlnLeuArgAspLeuAspIlePheLeuSerIleValAspMetAsnGluGlyAlaLeu 450
 QY 1379 ATGCTCCACACTTCAACTCAAGGCCATGTTATCGTCTGCTCAACAAAGAGAACTGGA 1438
 DB 451 LeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGlyAspAla 470
 QY 1439 AACCTTGAATCTCGCTGTGAAGAAAGACACACAGAGGGGACGGCGGAGAGAG 1498
 DB 471 AsnIleGluLeuValGlyLeu---LysGluGlnGlnGlnGlnGlnGlnGlnGlnGln 489
 QY 1499 GAGGACGAAGACGAAGAGAGGGAAGTAAACAGAGAGGTGCTAGGTACACAGCAGG 1558
 DB 490 Pro-----LeuGluValArgLysTyrArgAlaGlu 499
 QY 1559 TTGAAGAGAGCGATGTGTTTCATCATGCCAGCTCATCCAGTAGCCATCAACGCTTCC 1618
 DB 500 LeuSerGluGlnAspIlePheValIleProAlaGlyTyrProValValValAsnAlaThr 519
 QY 1619 TCCGACCTCATCTGCTTGGCTTCGATCAACGCTGAACCAACCAACAGATCTTCTT 1678
 DB 520 SerAsnLeuAsnPhePheAlaIleGlyIleAsnAlaGluAsnAsnGlnArgAsnPheLeu 539
 QY 1679 GCAGGTGATAAGGACAAATGTGATAGACCAAGATAGAGAAAGCAAGCAAGATTTACATTC 1738
 DB 540 AlaGlySerGlnAspAsnValIleSerGlnIleProSerGlnValGlnGlnLeuAlaPhe 559
 QY 1739 CCTGGTCCGGTGAACAAAGTTTGAGAGCTCATCAAAAAACAGAGAAATCTCACTTGTG 1798
 DB 560 ProGlySerAlaGlnAlaValGluLysLeuLeuLysAsnGlnArgGluSerTyrPheVal 579
 QY 1799 AGTGTCTCTCTCAATCTCAATCTCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1858
 DB 580 AspAlaGlnProLysLys----- 585


```
QY 1505 GAAGACCAAGAGGAGGAGTAACAGAGAGGTGCGTAGCTACACAGCGAGGTGTAAG 1564
||||| :|||
450 GluAspArgLysGluArg-----AsnAsnGluValGlnArgTyrGluAlaArgLeuSer 467
||||| :|||
QY 1565 GAAGGCGATGTTCATCATGCGAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAA 1624
||||| :|||
Db 468 ProGlyAspValIleIleProAlaGlyHisProValAlaIleSerAlaSerAsn 487
||||| :|||
QY 1625 CTCCTCTCTGCTGGCTTCATCAACGCTGTAACCAACCAACAGATCTCTTCAGGT 1684
||||| :|||
Db 488 LeuAsnLeuLeuGlyPheGlyIleAsnAlaLysAsnGlnArgAsnPheLeuSerGly 507
||||| :|||
QY 1685 GATAAGCAAAATGTGATACACAGATAGAGCAAGCAAGCAAGATTTAGCATTTCCCTGGG 1744
||||| :|||
Db 508 SerAspAspAsnValIleSerGlnIleGluAsnProValLysGluLeuThrPheProGly 527
||||| :|||
QY 1745 TCGGGTGAACAAGTTGAGAGCTCATCAAAAACCAAGAGGAATCTCACTTTGTGAGTGT 1804
||||| :|||
Db 528 SerSerGlnGluValAsnArgLeuIleLysAsnGlnLysGlnSerHisPheAlaSerAla 547
||||| :|||
QY 1805 CGTCTCAATCTCAATCTCAATCTCGTCTCGTCTCTGAGAAAGAGTCTCTCGAAGAG 1864
||||| :|||
Db 548 GluPro-----Glu 550
QY 1865 GATCAAGAGGAGGAAACCAAGGAGGAGGTCCTCTCTTCAATTTTGAAGGCTTTT 1924
||||| :|||
Db 551 GlnLysGluGluSerGlnArgLysArgSerProLeuSerSerValLeuAspSerPhe 570

RESULT 4
B24810
beta-conglycinin alpha chain precursor - soybean
C:Species: Glycine max (soybean)
C:Date: 07-Mar-1988 #sequence_revision 30-Jun-1991 #text_change 09-Jul-2004
C:Accession: B24810; S16337
R:Doyle, J.J.; Schuler, M.A.; Godette, W.D.; Zenger, V.; Beachy, R.N.; Slightom, J.L.
J. Biol. Chem. 261, 9228-9238, 1986
A:Title: The glycosylated seed storage proteins of Glycine max and Phaseolus vulgaris. S
A:Reference number: A24810; MUID:86250867; PMID:3013679
A:Accession: B24810
A:Molecule type: DNA
A:Residues: 1-639 <DOY>
A:Cross-references: UNIPROT:P11827; UNIPARC:UPI000012B569; GB:M13759; NID:g169928; PIDN:
A>Note: the authors translated the codon GGT for residue 352 as Glu
R:Schuler, M.A.; Schmitt, E.S.; Beachy, R.N.
Nucleic Acids Res. 10, 8225-8244, 1982
A:Title: Closely related families of genes code for the alpha and alpha' subunits of the
A:Reference number: S16337; MUID:83143288; PMID:6298713
A:Accession: S16337
A:Molecule type: DNA
A:Residues: 361-639 <SCH>
A:Cross-references: UNIPARC:UPI0000177DF9
C:Genetics:
A:Introns: 278/1; 355/3; 382/3; 481/3; 575/1
C:Superfamily: glycinin
C:Keywords: seed, storage protein

Alignment Scores:
Pred. No.: 2,07e-79 Length: 639
Score: 1251.00 Matches: 283
Percent Similarity: 58.7% Conservative: 120
Best Local Similarity: 41.3% Mismatches: 173
Query Match: 34.9% Indels: 110
DB: 2 Gaps: 21

US-10-728-323-1 (1-2032) x B24810 (1-639)
QY 50 ATGAGAGGAGGTTTCTCCACTGATGCTGTGCTAGGATCCTGTCTCTGCTTCAGTT 109
Db 2 MetArgAlaArgPhePro-----LeuLeuLeuLeuGlyValValPheLeuAlaSerVal 19
QY 110 TCTGCAACGCATGCCCAAGTCATCACCTTACCAGAGAAACAGAGAACCCCTCGCCCCAG 169
||||| :|||
||||| :|||
```

```
Db 20 SerValSerPheGlyIleAla-----TyrTrpGluLys-----GlnAsnProSerHisAsn 36
QY 170 AGGTCCTCCAGAGTTGTCAACAGGAACCGATGACTTTGAAGCAAAAGCATGCCAGTCT 229
||||| :|||
Db 37 LysCysLeuArgSerCysAsnSerGluLysAspSerTyrArgAsnGlnAlaCysHisAla 56
||||| :|||
QY 230 CGCTGCACC-----AAGCTCGAGTATGATCCTCTCTGT- 271
||||| :|||
Db 57 ArgCysAsnLeuLeuLysValGluGluGluGluCysGluGluGlyGlnIleProArg 76
||||| :|||
QY 272 CTTCTGA-----GGACACACTGGCACCAACCAACGCTTCCCTCCAGGGAGGGACACGT 328
||||| :|||
Db 77 ProArgProGlnHisProGluArgGluArgGlnGlnHisGlyGlyLysGluAspGlu 96
||||| :|||
QY 329 GSCCCCAACCCGGAGACTACGATGATGACGCG-----CGTCAACCCCGAAGAGAGAA--- 382
||||| :|||
Db 97 GlyGluGlnProArgPropheProArgProGlnProHisGlnGluGlu 116
||||| :|||
QY 383 -----GGAGCCCGATGG 394
||||| :|||
Db 117 HisGluGlnLysGluGluHisGluTrpHisArgLysGluGluLysHisGlyGlyLysGly 136
||||| :|||
QY 395 GGA-----CCAGCTGGACCGAGGAGCGTGA 421
||||| :|||
Db 137 SerGluGluGluGlnAspGluArgGluHisProArgProHisGlnProHisGlnLysGlu 156
||||| :|||
QY 422 AGAGAAGAA---GACTGGAGACAACCAAGAGAGATTGGAGGCGCAAGTCAATCAGCAG 478
||||| :|||
Db 157 GluGluLysHisGluTrpGlnHisLysGlnLysHisGlnGlyLysGluSerGluGlu 176
||||| :|||
QY 479 CCACGGAAATAAGCCCGCAAGAGAGAGAAAGCAACAGAG---TGGGGAACACACAGGT 535
||||| :|||
Db 177 GluGluGluAspGlnAspGluGluGlnAspLysGluSerGlnGluSerGluGly 196
||||| :|||
QY 536 AGCCATGTGAGGAGAAACAATCTCGAAC-----AACCTTTCTACTTCCGTC 586
||||| :|||
Db 197 SerGluSerGlnArgGluProArgArgHisLysAsnLysAsnProPheHisPheAsnSer 216
||||| :|||
QY 587 AGGCGGTTTACGCCGCTAGCGGAACCAAAACGCTAGATCCGGTCTCTCAGAGGTTT 646
||||| :|||
Db 217 LysArgPheGlnThrLeuPheLysAsnGlnTyrGlyHisValArgValGluGlnArgPhe 236
||||| :|||
QY 647 GACCAAAAGTCAAGGCAGTTTTCAGAAATCTCCAGAAATCACCGTATTGTGCAGATCGAGCC 706
||||| :|||
Db 237 AsnLysArgSerGlnGlnLeuGlnAsnLeuArgAspTyrArgIleLeuGluPheAsnSer 256
||||| :|||
QY 707 AAACCTAACACTTTGTTCTTCCCAAGCAGCGTGTATGTGTATTAATCCTTTTATCCAG 766
||||| :|||
Db 257 LysProAsnThrLeuLeuLeuProHisHisAlaAspAlaAspTyrLeuIleValIleLeu 276
||||| :|||
QY 767 CNAAGGCAAGCCCGTACCGTAGCAATGCGCAATACAGAAAGAGCTTTAATCTTGAC 826
||||| :|||
Db 277 AsnGlyThrAlaIleLeuThrLeuValAsnAsnAspArgAspSerTyrAsnLeuGln 296
||||| :|||
QY 827 GAGGCGCATGCACTCAGAAATCCATCCGTTTTCATTTCTCATCTTGAAC----- 877
||||| :|||
Db 297 SerGlyAspAlaLeuArgValProAlaGlyThrThrPheTyrValValAsnProAspAsn 316
||||| :|||
QY 878 -----CGCATGACAAAC----- 889
||||| :|||
Db 317 AspGluAsnLeuArgMetIleAlaGlyThrThrPheTyrValValAsnProAspAsnAsp 336
||||| :|||
QY 890 CAGAACCTCAGAGTAGCTAAATCTCCATGCCCGTTTAAACACACCCCGGAGTTTCCAGGAT 949
||||| :|||
Db 337 GluAsnLeuArgMetIleThrLeuAlaIleProValAsnLysProGlyArgPheGluSer 356
||||| :|||
QY 950 TTCTTCCCGGAGCAGCGACCAATCATCTACTTGCAGGGCTTTCAGCAGGAATACG 1009
||||| :|||
Db 357 PhePheLeuSerSerThrGlnAlaGlnSerTyrLeuGlnGlyPheSerLysAsnIle 376
||||| :|||
QY 1010 TTGAGGCGGCTTCAATCGCGGAATTCATAGATACGAGGAGGCTGCTGTAGAGAGAT 1069
||||| :|||
Db 377 LeuGluAlaSerTyrAspThrLysPheGluGluIleAsnLysValLeuPheGlyArgGlu 396
||||| :|||
```



```
QY 1382 CTCCACACTTCAACTCAAGGCATGGTTATCGTCGTCTCAACAAAGAACTGGAAC 1441
Db 300 LeuProHisTyrAsnSerArgAlaIleValIleValThrValAsnGluGlyValGlyAsp 319
QY 1442 CTTGAACTCGTCTGTAAGAAAGACCAACACAGAGGGGACGGCGGAAGAAGAGGAG 1501
Db 320 PheGluLeuValGlyGlnArgAsnGluAenGlnGlnGlnGlnArgLysGluAspGlu 339
QY 1502 GAGCAAGACGAAGAAGAGAGGGAAGTAACAGAGAGTGCCTAGGTACACAGCGGTTG 1561
Db 340 GluGluGluGlnGlnGluGluGluIleAsnLysGlnValGlnAsnTyrLysAlaLysLeu 359
QY 1562 AAGAAAGGCGATGTTTCATCATGCCAGAGCTCATCCAGTAGCCATCAACGCTTCTCTC 1621
Db 360 SerSerGlyAspValPheValIleProAlaGlyHisProValAlaValLysAlaSerSer 379
QY 1622 GAATCCATCTCGTTCGGTATCAACGCTGAAAGCAACACAGAGTCTCTCTGCA 1681
Db 380 AsnLeuAspLeuLeuGlyPheGlyIleAsnAlaGluAsnGlnArgAsnPheLeuAla 399
QY 1682 GGTGATAAGGACAATGTGATAGACCAAGCAAGCAAGGAGGATTTAGCATTTCCCT 1741
Db 400 GlyAspGluAspAsnValIleSerGlnIleGlnArgProValLysGluLeuAlaPhePro 419
QY 1742 GGGTCGGGTGAACAAGTTGAGAGCTCATCAAAACACAGAGAAATCTCATTGTGAGT 1801
Db 420 GlySerAlaGlnGluValAspArgIleLeuGluAsnGlnLysGlnSerHisPheAlaAsp 439
QY 1802 GCTCGTCTCAATCTCAATCTCCGTCCTCTGAGNAGAGTCTCTCGAGAA 1861
Db 440 AlaGlnProGlnGlnArg-----GluArgGlySerArgGluThr 452
QY 1862 GAGGATCAA 1870
Db 453 ArgAspArg 455

RESULT 6
S00567
vicilin precursor (clone pDUB9) - garden pea (fragment)
C:Species: Pisum sativum [garden pea]
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 05-Oct-2004
C:Accession: S00567
R:Watson, M.D.; Lambert, N.; Delauney, A.; Yarwood, J.N.; Croy, R.R.D.; Gatehouse, J.A.;
Biochem. J. 251, 857-864, 1988
A:Title: Isolation and expression of a pea vicilin cDNA in the yeast Saccharomyces cerev
A:Reference number: S00567; MUID:88326226; PMID:3046604
A:Accession: S00567
A:Molecule type: mRNA
A:Residues: 1-433 <WAT>
A:Cross-references: UNIPROT:PI3918; UNIPARC:UPI000016DF3C; EMBL:Y00722; NID:g20915; PIDN
F:1-2/Domain: signal sequence (fragment) #status predicted <SIG>
F:3-433/Product: vicilin #status predicted <MAT>

Alignment Scores:
Pred. No.: 4,17e-74 Length: 433
Score: 1174.50 Matches: 221
Percent Similarity: 72.2% Conservative: 93
Best Local Similarity: 50.8% Mismatches: 108
Query Match: 32.8% Indels: 13
DB: 2 Gaps: 4

US-10-728-323-1 (1-2032) x S00567 (1-433)

QY 566 AACCTTTTCTACTTCCGTCAGGCGGTTTAGCACCCGCTACGGNACCAAAACGGTAGG 625
Db 8 AsnProPheIlePheLysSerAsnLysPheGlnThrLeuPheGluAsnGlnGlyHis 27
QY 626 ATCCGGGTCTCGAGAGGTTTGACCAAGGTCAAGCGATTTTCAGAAATCTCCAGAAATCAC 685
Db 28 IleArgLeuLeuGlnLysPheAspGlnArgSerLysIlePheGluAsnGlnAsnTyr 47
QY 686 CGTATTGTGCAGATCAGAGCCAAACCTAACTCTTGTGTTCTTCCCAAGCACGCTGATGCT 745
```

```
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
48 ArgLeuLeuGluTyr-LysSerLysProHisThrIlePheLeuProGlnHisThrAspAla 67
746 GATAACATCTCTGTTATCCAGCAAGGCAAGCCACCGTGACCGTAGCAAAATGCCAATAAC 805
68 AspTyrIleLeuValValLeuSerGlyLysAlaIleLeuThrValLeuLysProAspArg 87
806 AGAAAGAGCTTTAATCTTTGACGAGGGCCATGCACCTCAGAATCCCATCCCGTTTCATTTCC 865
88 ArgAsnSerPheAsnLeuGluArgGlyAspThrIleLysLeuProAlaGlyThrIleAla 107
866 TACATCTTTGAACCGCCATGACCAACAGAACCTCAGAGTAGCTAAATACTCCATGCCCGTT 925
108 TyrLeuValAsnArgAspAsnGluGluLeuArgValLeuAspLeuAlaIleProVal 127
926 AACACACCGCGCCAGTTTGAGGATTTCTCCCGCGGAGCAGCGCAGACCAATCATCTCTAC 985
128 AsnArgProGlyGlnLeuGlnSerPheLeuLeuSerGlyAsnGlnAsnGlnGlnAsnTyr 147
986 TTGACAGGCTTCAGCAGGAATACGTTGAGGCGCGCTTCAATTCGGAATTTCAATGAGATA 1045
148 LeuSerGlyPheSerLysAsnIleLeuGluAlaSerPheAsnThrAspTyrGluGluIle 167
1046 CGGAGGTGCTGTTTGAAGAGAAATCCAGAGGTGAGCAAGAGGAGCAGAGCGCAGAGCGCA 1105
168 GluLysValLeuLeuGluGluHisGluLysGluThrGlnHisArgArgSerLeuLys--- 186
1106 TGGAGTACTCGGAGTAGTGAGAAACAATGAAGGAGTGATGTCAAAGTGTCAAAGGAGCAC 1165
187 ---AspLysArgGlnGlnSerGlnGluAsnValIleValLysLeuSerArgGlyGln 205
1166 GTTGAAGAACTTTACTAAGCACGCTAAATCCGCTCTCAAAAGAAAGGCTCCGCAAGAGGGA 1225
206 IleGluLeuLeuSerLysAsnAlaLysSerThrSerLysLysSerValSerSerGluSer 225
1226 GATATCAACCAACCGCAATCACTTGAGAGAGCGGAGCGCGGATCTTTCTAAACACTTTGGG 1285
226 Glu-----ProPheAsnLeuArgSerArgGlyProIleTyrSerAsnGluPheGly 242
1286 AAGTTATTGAGTGTAAGCGACACAAGAGAACCCCGACTTCAGGACTTCGACACATGATG 1345
243 LysPhePheGluIleThrProGlu---LysAsnProGlnLeuGlnAspLeuAspIlePhe 261
1346 CTCACCTGTGTAGAGATCAAAAGAGAGGAGCTTTGATGCTCCCACTTCAACTCAAGGCGC 1405
262 ValAsnSerValGluIleLysGluSerLeuLeuLeuProHisTyrAsnSerArgAla 281
1406 ATGTTTATCGTCGTCTCAACAAAGAACTTGAACCTTGAACCTCGTGGCTGTAGAAAA 1465
282 IleValIleValThrValAsnGluGlyLysGlyAspPheGluLeuValGlyGlnArgAsn 301
1466 GAGCAACAACAGAGGGGACGGCGGAAGAAGAGGAGGAGCAAGACGAGAGAGGAGGGA 1525
302 GluAsnGlnGlnGlnArgLysGluAspGluGluGluGluGluGluGluGlu 321
1526 AGTAAACAGAGAGTGCCTAGGTACACAGCGAGGTTTGAAGGAGGCGATGTGTTCATCATG 1585
322 IleAsnLysGlnValGlnAsnTyrLysAlaLysLeuSerSerGlyAspValPheValIle 341
1586 CCAGCAGCTCATCCAGTAGCCATCAACGCTTCTCCGAACTCCATCTCTGCTGGCTTCGGT 1645
342 ProAlaGlyHisProValAlaLeuLysAlaSerSerAsnLeuAspLeuGlyPheGly 361
1646 ATCAACGCTGAAACAACACACAGATCTCTTTCAGGTGTAGTAAAGCAATCTGTATAGAC 1705
362 IleAsnAlaGluAsnAsnGlnArgAsnPheLeuAlaGlyAspGluAspAsnValIleSer 381
1706 CAGATAGAGAGACGACGAGGATTTAGCATTTCCCTGGGTCCGGTGAACAACAGTTGAGAG 1765
382 GlnValGlnArgProValLysGluLeuAlaPheProGlySerAlaGlnGluValAspArg 401
1766 CTCATCAAAAACAGAGGAATCTCACTTTGTGAGTGTCTGCTCTCAATCTCAATCTCAA 1825
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

Db 402 IleLeuGluAsnGlnLysGlnSerHisPheAlaAspAlaGlnProGlnGlnArg----- 419

QY 1826 TCTCGTCTCTCTGAGAAAGAGTCTCTGAGAAAGAGGATCAA 1870
|||||
Db 420 -----GluArgGlySerArgGluThrArgAspArg 429
|||||

RESULT 7
FWSYCB
beta-conglycinin beta chain - soybean
C;Species: Glycine max (soybean)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 05-Oct-2004
C;Accession: JQ0969
R;Harada, J.J.; Barker, S.J.; Goldberg, R.B.
plant Cell 1, 415-425, 1989
A;Title: Soybean beta-conglycinin genes are clustered in several DNA regions and are reg
A;Reference number: JQ0969; MUID:93005638; PMID:2562562
A;Accession: JQ0969
A;Molecule type: DNA
A;Residues: 1-439 <HAR>
A;Cross-references: UNIPROT:P25974; UNIPARC:UPI000012B55E; GB:S44893; NID:G256426; PIDN:
C;Comment: This protein accumulates during seed development and is hydrolyzed after germ
C;Genetics:
A;Gene: CG-4
A;Introns: 101/1; 159/3; 186/3; 281/3; 375/1
C;Superfamily: vicilin
C;Keywords: seed; storage protein

Alignment Scores:
Pred. No.: 8.18e-73 Length: 439
Score: 1156.00 Matches: 237
Percent Similarity: 70.0% Conservatives: 87
Best Local Similarity: 51.2% Mismatches: 85
Query Match: 32.2% Indels: 54
DB: 1 Gaps: 10

US-10-728-323-1 (1-2032) x FWSYCB (1-439)

QY 542 GTGAGGAAGAACATCTCGGAACACACCCTTTCTACTTC---CCGTCAAGCGGTTTACG 598
|||||
Db 26 ValArgGluAsp-----GluAsnAsnProPheTyrPheArgSerAsnSerPheGln 43
|||||

QY 599 ACCCGTCTACCGGGAACCAAAACGGTAGGATCCGGTCTCTGAGAGGTTTGACCAAGGTCA 658
|||||
Db 44 ThrLeuPheGluAsnGlnAsnValArgIleArgLeuLeuGlnArgPheAsnLysArgSer 63
|||||

QY 659 AGGAGTTTCTGAGATCTCCAGATCCAGTATTTGTGAGATCCAGGCCAACCTAACACT 718
|||||
Db 64 ProGlnLeuGluAsnLeuArgAspTyrArgIleValGlnPheGlnSerLysProAsnThr 83
|||||

QY 719 CTGTTCTTCCCAAGCAGCGTGTGATGATAACATCTCTGTTATCCAGCAAGGCGAAGCC 778
|||||
Db 84 IleLeuLeuProHisHisAlaAspAlaAspPheLeuLeuPheValLeuSerGlyArgAla 103
|||||

QY 779 ACCGTGACCGTAGCAATGCGAATAACAGAAAGAGCTTTAATCTTGACGAGGCGCATGCA 838
|||||
Db 104 IleLeuThrLeuValAsnAsnAspArgAspSerTyrAsnLeuHisProGlyAspAla 123
|||||

QY 839 CTGAGATCCATCCGTTTCTATCTTCTATCTTGAACCGCCATGACCAACCAACCTC 898
|||||
Db 124 GlnArgIleProAlaGlyThrThrTyrTyrLeuValAsnProHisAspHisGlnAsnLeu 143
|||||

QY 899 AGAGTAGCTAAATCTCCATGCCCGTTAAACACACCGCGCGAGTTTGAGGATTTCTCCCG 958
|||||
Db 144 LysIleLeuLysLeuAlaIleProValAsnLysProGlyArgTyrAspAspPhePheLeu 163
|||||

QY 959 GCGAGCAGCCGAGCAACATCTCTACTTCAGGGCTTCAGCAAGGAATACGTTGGAGGCC 1018
|||||
Db 164 SerSerThrGlnAlaGlnGlnSerTyrLeuGlnGlyPheSerHisAsnIleLeuGluThr 183
|||||

QY 1019 GCCTTCAATCGGGAATTCATGATACGAGGCGTCTGTAGNAGAGAATCGAGGAGGT 1078
|||||
Db 184 SerPheHisSerGluPheGluGluIleAsnArgValLeuPhe-----Gly 198
|||||

QY 1079 GAGCAAGAGGAGAGAGCGGCGAGGCGATGAGTACTCGGAGTAGTGAGAACATGAAGGA 1138
|||||
Db 199 GluGluGluGluGlnArgGlnGln-----GluGly 208
|||||

QY 1139 GTGATAGTCAAAGTGTCAAAGAGGACGTTGAGAACTTACTAAGCACGCTAAATCCGTC 1198
|||||
Db 209 ValIleValGluLeuSerLysGluGlnIleArgGlnLeuSerArgArgAlaLysSerSer 228
|||||

QY 1199 TCAGAGAAA-----GGCTCCGAAGAAGAGGGAGATATCACCAACCAATCAACTTGAGA 1252
|||||
Db 229 SerArgLysThrIleSerSerGluAspGlu-----ProPheAsnLeuArg 243
|||||

QY 1253 GAAGGCGAGCCGATCTTTCTAACAACTTTGGGAACTTTAGAGTGAAGCCAGACAAG 1312
|||||
Db 244 SerArgAsnProIleTyrSerAsnAsnPheGlyLysPhePheGluIleThrProGlu--- 262
|||||

QY 1313 AAGAACCCCGAGCTTCAGACCTGGACATGATGCTCACCTGTGTAGAGATCAAAGAGGA 1372
|||||
Db 263 LysAsnProGlnLeuArgAspLeuAspIlePheLeuSerSerValAspIleAsnGluGly 282
|||||

QY 1373 GCTTTGATGCTCCACACATCTCAACTCAAAGGCCATGTTATCGTCTCGTCAACAAGGA 1432
|||||
Db 283 AlaLeuLeuLeuProHisPheAsnSerLysAlaIleValIleLeuValIleAsnGluGly 302
|||||

QY 1433 ACTGGAACCTTGAACCTCGTGTGTAAGAAAAGACAAACAGAGGGGACGCGGGAA 1492
|||||
Db 303 AspAlaAsnIleGluLeuValGlyIle---LysGluGlnGlnLysGlnLysGlnGlu 321
|||||

QY 1493 GAAGAGGAGGACGAAGCAAGAGAGGAGGAAGTAACAGAGGAGTGTGAGTGTACACA 1552
|||||
Db 322 GluGluPro-----LeuGluValGlnArgTyrArg 331
|||||

QY 1553 GCGAGGTTCAAGGAAGCGATGTTTCATCATGCCAGCAGCTCATCCAGTAGCCATCAAC 1612
|||||
Db 332 AlaGluLeuSerGluAspValPheValIleProAlaIleTyrProPheValValAsn 351
|||||

QY 1613 GCTTCTCTCGAACTCATCTGCTTGGCTTCAACGCTGATCAACGCTGAAAAACACACAGAATC 1672
|||||
Db 352 AlaThrSerAsnLeuAsnPheLeuAlaPheGlyIleAsnAlaGluAsnAsnGlnArgAsn 371
|||||

QY 1673 TTCCTTGCAGGTGATTAAGACCAATGTGATAGACCATGATAGAGAGCAAGCAAGGATTTA 1732
|||||
Db 372 PheLeuAlaGlyLysAspAsnValValArgGlnIleGluArgGlnValGlnGluLeu 391
|||||

QY 1733 GCATTCCTCGGTGCGGTGAACAAGTTGAGAACTCATCAAAAAACAGAGGAATCTCAC 1792
|||||
Db 392 AlaPheProGlySerAlaGlnAspValGluArgLeuLeuLysGlnArgGluSerTyr 411
|||||

QY 1793 TTTGTGAGTGCTCGTCTCAATCTCAATCTCAATCTCCGCTCTCTCTGAGAAAGAGTCT 1852
|||||
Db 412 PheValAspAlaGlnProGln----- 418
|||||

QY 1853 CTGAGAAAGAGGATCAAGAGGAGGAGAAACCAAGGAGGAGGTCCTCCTTTCAATT 1912
|||||
Db 419 -----GlnLysGluGluGlySerLysGlyArgLysGlyProPheProSerIle 434
|||||

QY 1913 TTGAAGGCT 1921
|||||

Db 435 LeuGlyAla 437
|||||

RESULT 8
S35757
vicilin, 47K - garden pea
C;Species: Pisum sativum (garden pea)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Oct-2004
C;Accession: S35757
R;Bown, D.
submitted to the EMBL Data Library, June 1992
A;Reference number: S35747
A;Accession: S35757
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-438 <BOW>

Db 295 LeuLeuProAsnTyrAsnSerArgAlaIleValIleValThrValThrGluGlyLysGly 314

Qy 794 AATGGCAATAACAGAAAGAGCTTTAATCTTGACGAGGGCCATGCACTCAGAAATCCCATCC 853


```
QY 974 CAATCATCTCTTGCAGGCGCTTCAGCAGGAATACGTTGAGGCGCCCTTCAATGCGGAA 1033
Db 170 GlnProSerIleLeuSerGlyPheSerLysAsnIleLeuGluAlaSerPheAsnThrArg 189
QY 1034 TTCAAATGATACGAGGGTCTGTTAGAGAAATGCGAGGAGTACGAGGAGGAGAG 1093
Db 190 TyrLysGluIleGluLysValLeuLeuGluGluHisGlyLysGluLysTyrHisArgArg 209
QY 1094 GGGCAGAGCGGATGAGTACTCGGAGTAGTGAGACAATGAGGAGGTAGTCAAGTGG 1153
Db 210 GlyLeuLysAspArgGlnArgGlyGlnGluLysAsn-----ValIleValLysIle 227
QY 1154 TCAAGGAGCAGCTTGAAGAACTTACTTAAGCAGCGTAAATCCGCTCAAGAAAGGCTCC 1213
Db 228 SerArgLysGlnIleGluLeuAsnLysAsnAlaLysSerSerLysLysSerThr 247
QY 1214 GAAGAAGAGGAGATATACCAACCCCAATCACTTGAGAGAGGCGAGCCGATCTTCT 1273
Db 248 SerSerGluSerGlu-----ProPheAsnLeuArgSerArgGluProIleTyrSer 264
QY 1274 AACAACCTTTGGGAAGTTATTGTAGGTGAGCCAGACAAGAAAGACCCAGCTTCAGGAC 1333
Db 265 AsnLysPheGlyLysPheGluIleThrPro---LysArgAsnProGlnLeuGlnAsp 283
QY 1334 CTGCACATGATGCTACCTGTGTAGAGATCAAAAGAGGAGCTTTGTATGCTCCACACTTC 1393
Db 284 LeuAsnIlePheValAsnTyrValGluIleAsnGluGlySerLeuLeuLeuProHisTyr 303
QY 1394 AACTCAAGGCCATGTTATCGTCGTCGTCACAAAGGAACTGGAACCTTGAACTCGTG 1453
Db 304 AsnSerArgAlaIleValIleValThrValAsnGluGlyLysGlyAspPheGluLeuVal 323
QY 1454 GCTGTAAAGAAAAGACAACAGAGGGGACGCGGAGAGAGAGGAGGAGGAGGAGAA 1513
Db 324 GlyGlnArgAsnGlnAsnGlnGlnGlyLeuArgGluGluTyrAspGluGluLysGluGln 343
QY 1514 GAAGAGGAGGAAAGTAACAGAGAGGTGCGTAGTACACAGCGAGGTGGAAGGAGCGAT 1573
Db 344 GlyGluGluGluIleArgLysGlnValGlnAsnTyrLysAlaLysLeuSerProGlyAsp 363
QY 1574 GTGTTTCATGTCAGCAGCTCATCCAGTAGCCATCAACGCTTCCTCCGAACTCCATCTG 1633
Db 364 ValLeuValIleProAlaGlyTyrProValAlaIleLysAlaSerSerAsnLeuAsnLeu 383
QY 1634 CTGCGTTCGATCAACGCTGAAACCAACACAGAAATCTTCTTCAGGTGATAAGGAC 1693
Db 384 ValGlyPheGlyIleAsnAlaGluAsnAsnGlnArgTyrPheLeuAlaGlyGluGluAsp 403
QY 1694 AATGTGATAGACAGATAGAGAAGCAAGCGAAGGATTTAGCATTCCTCGGTCGGTGAA 1753
Db 404 AsnValIleSerGlnIleHisLysProValLysGluLeuAlaPheProGlySerAlaGln 423
QY 1754 CAAGTTGAGAGCTCATCAAAAACCGAAGGAATCTCACTTTGTAGTGTCTGCTCTCAA 1813
Db 424 GluValAspThrLeuLeuGluAsnGlnLysGlnSerHisPheAlaAsnAlaGlnProArg 443
QY 1814 TCTCAATCTCAATCTCGTCTCTCTGAGAAAGAGTCTCTCGAGAAAGAGGATCAAGAG 1873
Db 444 -----GluArgGlu 446
QY 1874 GAGGAAACCAAGAGGAGGAGGTCCTCTCTTCAATTTGAAGCTTTT 1924
Db 447 ArgGlySerGlnGluIleLysAspHisLeuTyrSerIleLeuGlySerPhe 463

RESULT 11
FWPMVB
vicilin B precursor - garden pea (fragment)
C:Species: Pisum sativum (garden pea)
C:Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 05-Oct-2004
C:Accession: A03344
R:lycett, G.W.; Delauney, A.J.; Gatehouse, J.A.; Gilroy, J.; Croy, R.R.D.; Boulter, D.
Nucleic Acids Res. 11, 2367-2380, 1983
A:Title: The vicilin gene family of pea (Pisum sativum L.): a complete cDNA coding sequ
```

```
A:Reference number: A93462; MUID:83220791; PMID:6687941
A:Accession: A03344
A:Molecule type: mRNA
A:Residues: 1-410 <LYC>
A:Cross-references: UNIPROT:P02854; UNIPARC:UPI0000138276
A:Experimental source: cv. Feltham First, clones pDUB7 and pDUB4
A:Note: parts of this sequence, including the amino end of the mature protein, were det.
C:Comment: The gene that codes for this protein is part of a multigene family coding fo.
C:Superfamily: vicilin
C:Keywords: glycoprotein
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-410/Product: vicilin type B (fragment) #status predicted <MAT>
F:321-322/Cleavage site: Asn-Asp (unidentified proteinase) #status experimental
F:359/Binding site: carboxylate (Asn) (covalent) #status experimental

Alignment Scores:
Pred. No.: 1,12e-66 Length: 410
Score: 1068.00 Matches: 209
Percent Similarity: 71.6% Conservative: 84
Best Local Similarity: 51.1% Mismatches: 104
Query Match: 29.8% Indels: 12
DB: 1 Gaps: 5

US-10-728-323-1 (1-2032) x FWPMVB (1-410)
QY 542 GTGAGGGAAGAAACATCTCGAACAACCCCTTTCTACTTCCGTCAGGCGGTTTACACC 601
Db 13 ValSerSerArgSerAspGlnGluAsnProPheIlePheLysSerAsnArgPheGlnThr 32
QY 602 CGCTACGGGAACCAAAACGGTAGGATCCGGGTCCTGCAGAGGTTTGCACCAAGGTCAGG 661
Db 33 LeuTyrGluAsnGluAsnGlyHisIleArgLeuLeuGlnLysPheAspLysArgSerLys 52
QY 662 CAGTTTCAGAAATCTCAGAAATCACCGTATTGTGCAGATCGAGGCCAAACCTACACTCTT 721
Db 53 IlePheGluAsnLeuGlnAsnTyrArgLeuLeuGluTyrLysSerLysProHisThrLeu 72
QY 722 GTTCTTCCCAAGCAGCTGATCTGATAACATCTTGTATTCATCCAGCAAGGCGCAAGCCACC 781
Db 73 PheLeuProGlnTyrThrAspAlaAspPheIleLeuValValLeuSerGlyLysAlaThr 92
QY 782 GTGACGCTAGCAAAATGGCAATTAACAGAAAGAGCTTTTAATCTTGACGAGGCGCATGACATC 841
Db 93 LeuThrValLeuLysSerAsnAspArgAsnSerPheAsnLeuGluArgGlyAspAlaIle 112
QY 842 AGAATCCCATCCGGTTTCATTTCTTACATCTTGAACCGCATGACAAACAGAACCTCAGA 901
Db 113 LysLeuProAlaGlySerIleAlaTyrPheAlaAsnArgAspAspAsnGluGluProArg 132
QY 902 GTAGCTAAATCTCCATGCCCCGTTTAAACACCCCGGCAGTTTTCAGGATTTCTTCCCGCG 961
Db 133 ValLeuAspLeuAlaIleProValAsnLysProGlyGlnLeuGlnSerPheLeuLeuSer 152
QY 962 ACAGCCGAGACCAATCATCTTCTGAGGCTTCAGAGGCTTCAGCAGGAATACGTTGAGGCGGCC 1021
Db 153 GlyThrGlnAsnGlnLysSerSerLysSerGlyPheSerLysAsnIleLeuGluAlaIle 172
QY 1022 TTCAATCGGAATTCAAATGAGATACGAGGCTGTCTTAGAAGAGAAATCAGGAGGTGAG 1081
Db 173 PheAsnThrAsnTyrGluIleGluIleGluLysValLeuLeuGluGlnGln-----Glu 189
QY 1082 CAAGAGGAGAGGCGCAGGCGATGGAGTACTCGGAGTAGTGAG---AACATGAAGGA 1138
Db 190 GlnGluProGlnHisArgSerSerLeuLysAspArgArgGlnGlnIleAsnGluGluAsn 209
QY 1139 GTGATAGTCAAAAGTGTCAAGGAGCAGCTTGAAGAACTTACTTAAGCAGCACCTAAATCCGTC 1198
Db 210 ValIleValLysValSerArgAspGlnIleGluLeuSerLysAsnAlaLysSerSer 229
QY 1199 TCAAGAGAGGCTCCGAGAAAGAGGAGATATACCAACCCCAATCAACTCACTTGAGAGAGGC 1258
Db 230 SerLysLysSerValSerSerGlu-----SerGlyProPheAsnLeuArgSerArg 246
```


Db	538	VallLeuGluAlaAlaLeuAsnIleProArgAspLysLeuGluArgIlePhe		554
QY	1067	AATGCAGGAGTGCAGCAAGGAGGAGCGGATCGGAGTACTCGGAGTACTGAG	1126	554
Db	555	-----LysGlnArgArgGluArgGly	561	561
QY	1127	AACAATCAAGAGTGTAGTCAAGTGTCAAGGAGCAGCAGTTCAGAACTTACTAAGCAC	1186	1186
Db	562	-----GlyLysIleValArgAlaSerGlnGlnLeuArgAlaLeuSerGlnArg	578	578
QY	1187	GCTAATCCGTCTCAAGAAAGCTCCGAAAGAGGGAGATATACCAACCAATCAAC	1246	1246
Db	579	AlaThrSerVal---ArgLysGlySerArg-----GlyValArgAlaProIleLys	594	594
QY	1247	TTGAGAGAGGCGGAGCCGATCTTCTTCAACCTTTGGGAAGCTATTTCAGGTGAAGCCA	1306	1306
Db	595	LeuLysSerGlnThrProValThrAsnAsnGlnThrGlyGlnMetPheGluAlaCysPro	614	614
QY	1307	GACAAGAAGAACCCCGAGCTTCAGGACCTGGACATGATGTACCTGTGTAGAGATCAA	1366	1366
Db	615	Asp---GluPheProGlnLeuArgArgThrAspValAlaThrSerValValAspIleLys	633	633
QY	1367	GAAGAGCTTTGATGTCCACACTTCAACTCAAGCCATGTTATCGTGTGTCATC	1426	1426
Db	634	GlnGlyGlyMetMetValProHisPheAsnSerArgAlaThrTrpValValPheValSer	653	653
QY	1427	AAAGAACTGGAACCTTGAACCTCGTGGCT-----GTAAGAAAGAGCAACACAGAGG	1480	1480
Db	654	GluGlyAlaGlySerPheGluMetAlaCysProHisIleGlnSerSerGlnTrpGlnArg	673	673
QY	1481	GGACGCGGGAAGAAAG-----GAGGACCAAGACCAAGAGAGGAGGA	1525	1525
Db	674	GlyArgArgGluGluArgHisTrpArgArgGluGluGluGluArgGluGluArg	693	693
QY	1526	AGTAACAGAGAGTGGTGTACAGCAGAGGTTGAAGAGGCGATGTTTCATCATG	1585	1585
Db	694	SerGlyArg---PheGluArgValAlaGlyArgLeuSerGluGlyGlyValLeuValIle	712	712
QY	1586	CCAGCAGCTCATCCAGTACCAACAGCTTCC-----TCCGAACCTCCATCTGCTTGGC	1639	1639
Db	713	ProAlaGlyHisProIleAlaIleMetAlaSerProAsnGlnAsnLeuArgLeuValGly	732	732
QY	1640	TTCCGTATCAACCTGAAACACACAGAAATCTTCTTCAGGTGATTAAGGACATGTG	1699	1699
Db	733	PheGlyIleAsnAlaGluAsnAsnHisArgAsnPheLeuAlaGly---ArgGluAsnIle	751	751
QY	1700	ATAGACCATAGAGAGCAAGCAAGGATTTAGCATTTCCCTGGGTGGGTGAACAGTT	1759	1759
Db	752	MetAsnGluLeuAspArgGluAlaLysGluLeuAlaPheAsnValGluGlyLysGlnAla	771	771
QY	1760	GAGAAGCTCATCAAAACAGAGGAATCTCACTTTGTGAGTGCTCGTCTCAATCTCAA	1819	1819
Db	772	AspGluIlePheArgSerGlnArgGluSerPhePhe-----	783	783
QY	1820	TCTCAATCTCCGTGCTCTCTGAGAAAGAGTCTCTCGAAGAGGATCAAGAGGAGGAA	1879	1879
Db	784	ThrGluGlyProGluGlyArgArgArgSerThrGlu-----	796	796
QY	1880	AACCAAGGAGGAGGTTCACTTCTTTCAATTTTGAAG	1918	1918
Db	797	-----ArgSerProLeuLeuSerIleLeuLys	805	805

RESULT 15

S06398
alpha-globulin type A precursor - upland cotton
N;Alternate names: seed storage protein
C;Species: Gossypium hirsutum (upland cotton)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C;Accession: S06398
R;Chlan, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.
Plant Mol. Biol. 9, 533-546, 1987
A;Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX. Se

A;Reference number: S06398
A;Accession: S06398
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-605 <CHL>
A;Cross-references: UNIPROT:P09799; UNIPARC:UPI000011EF3D
C;Superfamily: signalin
F;1-24/Domain: glycinin
F;25-605/Product: alpha-globulin type A #status predicted <SIG>
Alignment Scores:
Pred. No.: 1.17e-50 Length: 605
Score: 839.00 Matches: 216
Percent Similarity: 50.9% Conservative: 108
Best Local Similarity: 33.9% Mismatches: 223
Query Match: 23.4% Indels: 90
DB: 2 Gaps: 21

US-10-728-323-1 (1-2032) x S06398 (1-605)

QY	65	TCCTCCACTGATGCTGTGCTAGGATCCTTGCTCGGCTTCAGTTCTGCAACGCATGCC	124			124
Db	6	SerValPheValIleValLeuPheSerLeuPheLeuSerPheGlyLeuLeuCysSerAla	25			25
QY	125	AAATCATCACCTTACCAGAGAAAAACAGAACCCCTCGCCCCAGAGG-----TGC	175			175
Db	26	LysAspPheProGlyArgArgSerGluAspAspPro---GlnGlnArgTyrGluAspCys	44			44
QY	176	CTCCAGAGTTGTCAACAGAACCG---GATGACTTGAAGCAAAAGGCATCGAGTCTCG	232			232
Db	45	ArgLysArgCysGlnLeuGluThrArgGlyGlnThrGlnGlnAspLysCysGluAspArg	64			64
QY	233	TGC---ACCAAGCTCGAG-----TATGATCCTCGTTGTGTCTAT	268			268
Db	65	SerGluThrGlnLeuLysGluGlnGlnArgAspGlyGluAspProGlnArgArgTyr	84			84
QY	269	---GATCTCTGAGGACACACTGGCACCAACCAACGTTCCCTCCAGG-----	316			316
Db	85	GlnAspCysArgGlnHisCysGlnGlnGluGluArgLeuArgProHisCysGluGln	104			104
QY	317	-----GAGCGGACACGTGGCCGCCACCGACAGCTGGACCG-----	409			409
Db	105	SerCysArgGluGlnTyrGluLysGlnGlnGlnGlnPro-----Asp	119			119
QY	359	CGCCCTCAACCCCGAAGAGAGGAGCGCGATGGACACAGCTGGACCG-----	409			409
Db	120	LysGlnPheLysGluCysGlnGlnArgCysGlnTrpGlnGluArgProGluArgLys	139			139
QY	410	-----AGGAGCGTGAAGAGAGAAAGACTGGAGACCAACCAAGA	448			448
Db	140	GlnGlnCysValLysGluCysArgGluGlnTyrGlnGluAspProTrpLysGlyGluArg	159			159
QY	449	GAAGAT---TGAGCGGACCAAGTCATCAGCAGCCACCGAAATAAGGCCCAAGAAAGA	505			505
Db	160	GluAsnLysTrpArgGluGluGluGlu-----GluSerAsp	172			172
QY	506	GAAGGAGAACAGAGTGGGGAAACACAGGTAGCCATGTGAGGGAAGAAACATCTCGGAAC	565			565
Db	173	GluGlyGluGln-----GlnArgAsn	180			180
QY	566	AACCTTTCTACTTCCCGTCAAGCGGTTTAGACCCGCTACGGAAACCAACCGTAGG	625			625
Db	181	AsnProTyrTyrPheHisArgSerPheGlnGluArgPheArgGluHisGlyAsn	200			200
QY	626	ATCCCGGCTCTGACAGGTTTACCAAGGTCAAGGCAGTTTCAGATCTCCAGATCAC	685			685
Db	201	PheArgValLeuGlnArgPheAlaAspLysHisLeuLeuArgGlyIleAsnGluPhe	220			220
QY	686	CGTATTGTGCAGATCGAGGCCAAACTTAACACTCTTGTCTTCCCAAGCAGCGTGTGT	745			745
Db	221	ArgGluAlaIleLeuGluAlaAsnProAsnThrPheValLeuProHisCysAspAla	240			240
QY	746	GATAACATCTTTGTTTATCCAGCAAGGCAAGCCCGTAGCAATGCAATAAC	805			805

Db 581 AspArgArgGlyPheAspGluArgArgGlySerAsnAsnProLeuSerPro 597

Search completed: May 15, 2006, 22:23:08
Job time : 109.003 secs

Db	241	GlulysileTyValValValThrAsnGlyArgGlyThrValThrPheValThrHisGluAsn	260
Qy	806	AGAAGAGCTTATTCITGACGAGGGCCATGCACCTCAGATCCCATCCGGTTTCATTTCC	865
Db	261	LysGluSerTyAsnValValProGlyValValValArgIleProAlaGlySerThrVal	280
Qy	866	TACATCTTGAACCCCATCATCAACACCAACTCAGAGTAGCTAAATCTCCATGCCGGTT	925
Db	281	TyrLeuAlaAsnGlnAspAsnArgGluLeuThrIleAlaValLeuHisArgProVal	300
Qy	926	AACACACCCGGCCAGTTTGAGGATTTCTTCCCGCGAGCAGCCGAGACCAATCATCTCTAC	985
Db	301	AsnAsnProGlyGlnPheGlnLysPhePheProAlaGlyGlnGluAsnProGlnSerTyr	320
Qy	986	TTGCAGGGCTTCACAGGAATACCTTTCGAGGCCCTTCAATCGGGAATTCATGAGATA	1045
Db	321	LeuArgIlePheSerArgGluIleLeuGluAlaValPheAsnThr	335
Qy	1046	CGGAGGGTCTGTTTGAAGAAATGCAAGAGGTGAGCAAGAGGAGAGCGCAGAGCGCA	1105
Db	336	ArgSerGluGlnLeuAspGluLeuProGlyGlyArgGlnSerHisArgArgGlnGln---	354
Qy	1106	TGGAGTACTCGGAGTAGTGAAACAATGAAGAGTAGTATGCAAGTGTCAAAGGAGCAC	1165
Db	355	-----GlyGlnGlyMetPheArgIysAlaSerGlnGluGln	366
Qy	1166	GTTGAGNACTTACTAAGCAGCGCTAAATCCGTCTCAAGAAGCGTCCGAGAGAGGGA	1225
Db	367	IleArgAlaLeuSerGlnGlyAlaThrSerProArgGlyLysGlySerGluGlyTyrAla	386
Qy	1226	GATATCACCAACCAATCAACTTCAGAGAAGCGGAGCCCGATCTTTCTAAACAATTTGGG	1285
Db	387	-----PheAsnLeuLeuSerGlnThrProArgTyrSerAsnGlnAsnGly	401
Qy	1286	AAGTTATTGTAGGTGAAGCCACAGAAAGAACCCCGAGTCTCAGACCTTGGACATGATG	1345
Db	402	ArgPheTyrGluAlaCysProArgAsnPheGlnGlnGlnLeuArgGluValAspSerSer	421
Qy	1346	CTCACCTGTGTAGAGATCAAGAGGAGCTTTGATGTCTCCACACTTCACTCAAGAGCC	1405
Db	422	ValValAlaPheGluIleAsnLysGlySerIlePheValProHisTyrAsnSerLysAla	441
Qy	1406	ATGGTTATCTGTCGTCGTCACAAAGGAACCTGAAACCTGAACCTCGTGGCT-----	1459
Db	442	ThrPheValValLeuValThrGluGlyAsnGlyHisValGluMetValCysProHisLeu	461
Qy	1460	AGAAAAGAACCAACACAGAGGGGACGGCGGAAGAGGAGGACGAAAGAGAGAGAG	1519
Db	462	SerArgGlnSerSerAspTyrSerArgGluGluGluGlnGluGluGlnGluVal	481
Qy	1520	GAGGGAAGTACAGAGAGGTCGTAGGTACACAGGAGGTGGAAGGAGCGGATGTCTTC	1579
Db	482	GluArgSerGlyGlnTyrLysArgValArgAlaGlnLeuSerThrGlyAsnLysPhe	501
Qy	1580	ATCATGCCAGAGCTCATCCAGTAGCCATCAACGCTTCC-----TCGAACTCCATCTG	1633
Db	502	ValValProAlaGlyHisProValThrPheValAlaSerGlnAsnGluAspLeuGlyLeu	521
Qy	1634	CTTGCTTCGGTATC---AAGCTGAAACCAACCAACCAATCTTCCCTTGAGGTGATAAG	1690
Db	522	LeuGlyPheGlyLeuTyrAsnGlyGlnAspAsnLysArgIlePheValAlaGly---Lys	540
Qy	1691	GACATGTGATAGACCATAGAGNAGCAAGCGAGGATTTAGCATTTCCCTGGGTCCGGT	1750
Db	541	ThrAsnAsnValArgGlnTrpAspArgGlnAlaLysGluLeuAlaPheGlyValGluSer	560
Qy	1751	GAAACAAGTTGAGAAGCTCATCAAAAACCGAG---AAGAACTCTCACTTTGTGAGTCTCGT	1807
Db	561	ArgLeuValAspGluValPheAsnAsnAsnProGlnGluSerTyrPheValSerGlyArg	580
Qy	1808	-----CCTCAATCTCAATCTCAATCTCCGTTCCTCTCTCT	1840

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 21:23:02 ; Search time 21.4781 Seconds

(without alignments)

4400.310 Million cell updates/sec

Title: US-10-728-323-2

Perfect score: 1281

Sequence: 1 gctcaccatactagtagcccc.....taaaagaatcatgttttggtt 717

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p model -DEV=xlp
-Q/abs/ABSSWEB spool/US10728323/runat_15052006_172130_22366/app_query.fasta.1
-DB=A_Geneseq -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US10728323 @CGN 1 1 605 @runat_15052006_172130_22366 -NCPU=6 -ICPU=3
-NO_MAP -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: Geneseq21.1*
- 2: Geneseqp1980s.*
- 3: Geneseqp1990s.*
- 4: Geneseqp2000s.*
- 5: Geneseqp2001s.*
- 6: Geneseqp2002s.*
- 7: Geneseqp2003as.*
- 8: Geneseqp2003bs.*
- 9: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	842	65.7	157	2 AAW24153	Aaw24153 Peanut al
2	842	65.7	157	2 AAW24164	Aaw24164 Peanut al
3	842	65.7	157	2 AAY15245	Aay15245 Peanut al
4	842	65.7	157	4 AAU04707	Aau04707 Anaphylac
5	842	65.7	157	6 ABUS2464	Abu52464 Peanut Ar
6	842	65.7	157	7 ADG27518	Adg27518 Peanut al
7	842	65.7	157	9 ADV97542	Adv97542 Peanut Ar
8	842	65.7	159	9 ADV97603	Adv97603 Peanut Ar
9	838.5	65.5	207	4 AAB82383	Aab82383 Peanut al

10	830	64.8	156	2 AAY40973	Aay40973 Ara h 2 p
11	827	64.6	157	9 ADV97553	Adv97553 Peanut Ar
12	819	63.9	156	8 ADO38314	Ado38314 Ara h 2 M
13	814	63.5	157	6 ABUS2576	Abu52576 Peanut Ar
14	812	63.4	156	8 ADM12096	Adm12096 Arachis h
15	812	63.4	157	6 ABUS2575	Abu52575 Peanut Ar
16	774	60.4	157	6 ABUS2577	Abu52577 Peanut Ar
17	771	60.2	166	2 AAY40968	Aay40968 Recombina
18	771	60.2	166	6 ABUS2482	Abu52482 Peanut Ar
19	771	60.2	166	7 ADG27536	Adg27536 T7/His-ta
20	699	54.6	166	3 AAB33600	Aab33600 Modified
21	699	54.6	166	3 AAU05035	Aau05035 Modified
22	688.5	53.7	167	4 AAU04710	Aau04710 Modified
23	243	19.0	158	2 AAW23419	Aaw23419 Soybean a
24	243	19.0	158	2 AAY05723	Aay05723 Soybean G
25	243	19.0	158	7 ADH89269	Adh89269 G. max 2S
26	243	19.0	158	8 ADG44004	Adg44004 G. max 2S
27	241.5	18.9	155	2 AAW23418	Aaw23418 Soybean a
28	241.5	18.9	155	7 ADH89267	Adh89267 G. max na
29	241.5	18.9	155	8 ADG44002	Adg44002 G. max 2S
30	229	17.9	158	2 AAW23420	Aaw23420 Chimeric
31	184.5	14.4	323	7 ADH89277	Adh89277 Sunflower
32	184.5	14.4	323	8 ADG44012	Adg44012 H. annuus
33	177.5	13.9	168	6 AAE36076	Aae36076 Flax Conl
34	156	12.2	28	2 AAW24194	Aaw24194 Peanut al
35	156	12.2	28	9 ADV97585	Adv97585 N-termina
36	156	12.2	158	2 AAW23586	Aaw23586 Mabinlin
37	151	11.8	158	2 AAW23588	Aaw23588 Mabinlin
38	151	11.8	295	7 ADH89275	Adh89275 Sunflower
39	151	11.8	295	8 ADG44010	Adg44010 H. annuus
40	149.5	11.7	167	8 ADS16393	Ads16393 Grape 2S
41	146.5	11.4	164	1 AAP91892	Aap91892 1kb fragm
42	146.5	11.4	164	1 AAP96144	Aap96144 Sequence
43	146.5	11.4	164	7 ADH89227	Adh89227 A. thalia
44	146.5	11.4	164	8 ADG43962	Adg43962 A. thalia
45	144	11.2	169	6 AAE36077	Aae36077 Flax Conl

ALIGNMENTS

RESULT 1

AAW24153

ID AAW24153 standard; protein; 157 AA.

XX

AC AAW24153;

XX

DT 17-OCT-2003 (revised)

DT 29-DEC-1997 (first entry)

XX

DE Peanut allergen Ara hII.

XX

KW Peanut; seed storage protein; allergen; allergy; hypersensitivity;

KW vaccine; anaphylactic shock; immunotherapy; therapy; monoclonal antibody;

KW ELISA; analysis; Ara hII.

XX

OS Arachis hypogaea; strain Florunner.

XX

PN WO9724139-A1.

XX

PD 10-JUL-1997.

XX

PF 23-SEP-1996; 96WO-US015222.

XX

PR 29-DEC-1995; 95US-0009455P.

XX

PR 04-MAR-1996; 96US-00610424.

XX

(UYAR-) UNIV ARKANSAS.

 PI | Burks AW, Helm RM, Cockrell G, Stanley JS, Bannon GA; | XX | DR | WPI; 1997-363453/33. | XX | DR | N-PSDB; AAT76615. | XX |

PT Peanut allergens Ara hi and Ara hII - used for vaccination and in two-
 PT site monoclonal antibody based ELISA.
 XX
 XX Claim 31; Page 198; 354pp; English.
 XX
 CC This polypeptide comprises major peanut allergen Ara hII. Its sequence
 CC was deduced from a cDNA clone (AA76615) isolated from peanut seed cDNA
 CC using a primer (see AAT76617) based on an isolated Ara hi peptide (see
 CC AAW24151). The sequence shows significant homology with the conglutin
 CC family of seed storage proteins of other legumes. The allergen is
 CC recognised by serum IgE from a large proportion of individuals with
 CC peanut hypersensitivity. Ara hII and Ara hi (see AAW24149-50) can be used
 CC to raise monoclonal antibodies which are used in a specific two-site MAB
 CC ELISA for the detection of Ara hi or Ara hII (claimed). IgE-binding Ara
 CC hII antigen epitopes (see AAW24188-93) may be used in vaccines to protect
 CC against allergic reactions to peanut allergens, e.g. anaphylactic shock.
 CC (Updated on 17-OCT-2003 to standardise OS field)
 XX
 XX Sequence 157 AA;

Alignment Scores:
 Pred. No.: 9,32e-90 Length: 157
 Score: 842.00 Matches: 157
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 65.7% Indels: 0
 DB: 2 Gaps: 0

US-10-728-323-2 (1-717) x AAW24153 (1-157)

Qy 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCGATCGGAGCAG 61
 Db 1 LeuThrIleuValAlaLeuAlaLeuPheLeuAlaAlaHisAlaSerAlaArgGln 20
 Qy 62 CAGTGGGAACCTCAAAGGAGACAGAGATGCCAGCCAGCTCGAGAGGGGCAACCTGAGG 121
 Db 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
 Qy 122 CCCTGCGAGCAACATCTTCATGCAAGATCCAACTGACGAGGATTCATATGAACGGGAC 181
 Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
 Qy 182 CCGTACAGCCCTAGTAGGATCCGTACAGCCCTAGTCCATATGTCGAGAGCGCTGGA 241
 Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
 Qy 242 TCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAACCAACCAAGG 301
 Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
 Qy 302 TGCATGTGCGAGGCAATTGCAACAGATCATGGAGAACCCAGAGCGGATGTTGCGAGGGAGG 361
 Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
 Qy 362 CNACAGAGACACAGTTCAAGAGGGAGCTCAGGNACTTGCTCAGACAGTCGGCCTTAGG 421
 Db 121 GlnGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
 Qy 422 GCACCACAGCGTTGCGACTTGGAGCTGCAAGTGGCGGCGGAGAGACAGATAC 472
 Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157

RESULT 2

AAW24164
 ID AAW24164 standard; protein; 157 AA.

XX AC AAW24164;

DT 17-OCT-2003 (revised)

DT 29-DEC-1997 (first entry)

XX Peanut allergen Ara hII.

DE
 XX

KW Peanut; seed storage protein; allergen; allergy; hypersensitivity;
 KW vaccine; anaphylactic shock; immunotherapy; therapy; monoclonal antibody;
 KW ELISA; analysis; Ara hII.

OS Arachis hypogaea; strain Florunner.

XX WO9724139-A1.

XX 10-JUL-1997.

XX 23-SEP-1996; 96WO-US015222.

XX 29-DEC-1995; 95US-0009455P.

XX 04-MAR-1996; 96US-00610424.

XX (UYAR-) UNIV ARKANSAS.

XX Burks AW, Helm RM, Cockrell G, Stanley JS, Bannon GA;

XX WPI; 1997-363453/33.

XX N-PSDB; AAT76614.

XX Peanut allergens Ara hi and Ara hII - used for vaccination and in two-
 PT site monoclonal antibody based ELISA.

XX Claim 31; Page 219; 354pp; English.

CC This polypeptide comprises major peanut allergen Ara hII. Its sequence
 CC was deduced from cDNA clone P38 (AAT76614), isolated from peanut seed
 CC cDNA using a primer (see AAT76617) based on an isolated Ara hi peptide
 CC (see AAW24151). The sequence shows significant homology with the
 CC conglutin family of seed storage proteins of other legumes. The allergen
 CC is recognised by serum IgE from a large proportion of individuals with
 CC peanut hypersensitivity. Ara hII and Ara hi (see AAW24149-50) can be used
 CC to raise monoclonal antibodies which are used in a specific two-site MAB
 CC ELISA for the detection of Ara hi or Ara hII (claimed). IgE-binding Ara
 CC hII antigen epitopes (see AAW24188-93) may be used in vaccines to protect
 CC against allergic reactions to peanut allergens, e.g. anaphylactic shock.
 CC (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 157 AA;

Alignment Scores:
 Pred. No.: 9,32e-90 Length: 157
 Score: 842.00 Matches: 157
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 65.7% Indels: 0
 DB: 2 Gaps: 0

US-10-728-323-2 (1-717) x AAW24164 (1-157)

Qy 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCGATCGGAGCAG 61

Db 1 LeuThrIleuValAlaLeuAlaLeuPheLeuAlaAlaHisAlaSerAlaArgGln 20

Qy 62 CAGTGGGAACCTCAAAGGAGACAGAGATGCCAGCCAGCTCGAGAGGGGCAACCTGAGG 121

Db 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40

Qy 122 CCCTGCGAGCAACATCTTCATGCAAGATCCAACTGACGAGGATTCATATGAACGGGAC 181

Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60

Qy 182 CCGTACAGCCCTAGTAGGATCCGTACAGCCCTAGTCCATATGTCGAGAGCGCTGGA 241

Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80

Qy 242 TCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAACCAACCAAGG 301

Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100

Qy 302 TGCATGTGCGAGGCAATTGCAACAGATCATGGAGAACCCAGAGCGGATGTTGCGAGGGAGG 361

CC Ara h 2. Ara h 2 is an anaphylactic antigen (A), which was used to design
 CC antigenic peptides having a reduced ability to bind immunoglobulin E
 CC (IgE) as compared with the intact (A), or having a sequence substantially
 CC identical to a portion of sequence of an antigen that includes at least
 CC one IgE binding site, where at least one IgE binding site of the peptide
 CC is altered. The antigenic peptides are used in a composition which is
 CC useful for reducing risk or severity of allergic reaction to an antigen.
 CC This is done by identifying an individual at risk of allergic reaction to
 CC an antigen by identifying prior display of allergic symptoms when exposed
 CC to the antigen, or a familial relationship with an individual who
 CC previously displayed allergic symptoms when exposed to the antigen.
 CC Following this an antigen-specific IgE present on one or more mast cells
 CC or basophils in the individual's serum is identified. The individual is
 CC then contacted with a peptide corresponding to a portion of the antigen,
 CC which is selected, formulated, and delivered so that binding of the
 CC peptide to antigen-specific IgE is reduced as compared with IgE binding
 CC of intact antigen. The composition is also useful for treating and
 CC preventing allergic reactions

XX Sequence 157 AA;

Alignment Scores:
 Pred. No.: 9.32e-90 Length: 157
 Score: 842.00 Matches: 157
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 65.7% Indels: 0
 DB: 4 Gaps: 0

US-10-728-323-2 (1-717) x AAU04707 (1-157)

Qy 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTGCTGCCACGGATCTGGAGGCAG 61
 Db 1 LeuThrileuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
 Qy 62 CAGTGGGAACCTCAAGGAGACAGAGATCCAGAGCCAGCTCCAGAGGGGGAACCTGAGG 121
 Db 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
 Qy 122 CCCTGCGAGCAACATCTCATGCGAAGATCCACGTGACGAGGATTCATATGAACGGAC 181
 Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
 Qy 182 CCGTACAGCCCTAGTCAAGATCCGTA CAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 241
 Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
 Qy 242 TCCTCTCAGCACCAAGAGAGTGTTGCAATGAGCTGAACGAGTTTGAGAACCAAGG 301
 Db 81 SerSerGlnHisGlnGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
 Qy 302 TGCATGTGCGAGGATTCGAACAGATCATGGAGAACCGAGCGCATAGTTGCGAGGGAGG 361
 Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
 Qy 362 CAACAGAGACACAGTTCCAGAGGGAGCTCAGGAACTTGCCCTCAACAGTGGCGCTTAGG 421
 Db 121 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
 Qy 422 GCACCAGACGCTTCGACGCTTGAGCTGCGAAGTGGCGGCGAGACAGATAC 472
 Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157

RESULT 5

ABU52464

ID ABU52464 standard; protein; 157 AA.

XX ABU52464;

XX 10-MAR-2003 (first entry)

XX Peanut Ara h2 protein.

XX

KW Peanut; allergy; Ara h1; Ara h2; Ara h3; IgE binding site;
 KW anaphylactic food allergen; antiallergenic; vaccine; wound healing.
 XX
 OS Arachis hypogaea.
 XX
 PN WO200274250-A2.
 XX
 PD 26-SEP-2002.
 XX
 XX 18-MAR-2002; 2002WO-US009108.
 XX
 PR 16-MAR-2001; 2001US-0276822P.
 PR 19-MAR-2002; 2002US-00276822.
 XX
 PA (PANA-) PANACEA PHARM.
 PI Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
 PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
 PI Rabin PA, Shin DS, Stanley JS;
 XX
 XX WPI; 2003-018765/01.
 DR N-PSDB; ABX70606.
 XX
 XX New modified anaphylactic food allergen, useful for preventing or
 PT treating allergic reactions associated with e.g. anaphylactic allergens.
 PT
 XX
 XX Claim 27; Fig 41; 300pp; English.
 PS
 XX
 CC The invention relates to a modified anaphylactic food allergen has an
 CC amino acid sequence that is substantially identical to that of natural
 CC anaphylactic food allergen, except for a cysteine residue that has been
 CC modified so that it cannot participate in the disulphide bond. The
 CC modification may also comprise mutation of the IgE binding sites to
 CC reduce allergenicity. Also included are: (1) a method of making a
 CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
 CC or for causing a site specific mutation in the modified anaphylactic food
 CC allergen; (3) a transgenic plant or animal expressing the modified
 CC anaphylactic food allergen; (4) a method of treating an individual by
 CC reducing the clinical response to a natural anaphylactic food allergen;
 CC and an isolated fragment of peanut allergen Ara h 1. The modified
 CC anaphylactic food allergen is useful for preventing or treating allergic
 CC reactions associated with any natural allergen such as food, insect,
 CC rubber or preferably anaphylactic allergens. It is also useful for
 CC treating wounds in mammals such as bovine, canine, feline, caprine,
 CC ovine, porcine, murine or equine species. The present sequence is a
 CC peanut allergen (e.g. Ara h1, h2 or h3)

XX Sequence 157 AA;

Alignment Scores:
 Pred. No.: 9.32e-90 Length: 157
 Score: 842.00 Matches: 157
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 65.7% Indels: 0
 DB: 6 Gaps: 0

US-10-728-323-2 (1-717) x ABU52464 (1-157)

Qy 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTGCTGCCACGGATCTGGAGGCAG 61
 Db 1 LeuThrileuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
 Qy 62 CAGTGGGAACCTCAAGGAGACAGAGATCCAGAGCCAGCTCCAGAGGGGGAACCTGAGG 121
 Db 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
 Qy 122 CCCTGCGAGCAACATCTCATGCGAAGATCCACGTGACGAGGATTCATATGAACGGAC 181
 Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
 Qy 182 CCGTACAGCCCTAGTCAAGATCCGTA CAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 241

Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
QY 242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAACACCAAGG 301
Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
QY 302 TGCATGTGCGAGGCATTGCAACAGATCATGGAGAACCAAGCGCATAGGTTGCGAGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CARACGAGCAACAGTTCAAGAGGGAGCTCAGAACTTGCCCTCAACAGTGGCGCCTTAGG 421
Db 121 GlnGlnGlnGlnGlnPheIlysrGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACCCAGCGTTTGGCACTTGGACGTGCGAAGTGGCGGAGAGACAGATAC 472
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157
RESULT 6
ADG27518
ID ADG27518 standard; protein; 157 AA.
XX
AC ADG27518;
XX
DT 26-FEB-2004 (first entry)
XX
DE Peanut allergen Ara h2.
XX
KW Peanut; plant; allergen; Ara h1; Ara h2; Ara h3; glycinin A2B1a; Jug n1;
KW antiallergic; vulnerability; anaphylactic food allergen; IgE; allergy; wound.
XX
OS Arachis hypogaea.
XX
PN US2003202980-A1.
XX
PD 30-OCT-2003.
XX
PF 18-MAR-2002; 2002US-00100303.
XX
PR 29-DEC-1995; 95US-0009455P.
PR 23-SEP-1996; 96US-00717933.
PR 31-JAN-1998; 98US-0073283P.
PR 13-FEB-1998; 98US-0074590P.
PR 13-FEB-1998; 98US-0074624P.
PR 13-FEB-1998; 98US-0074633P.
PR 29-JUN-1998; 98US-00106872.
PR 27-AUG-1998; 98US-00141220.
PR 13-NOV-1998; 98US-00191593.
PR 29-JAN-1999; 99US-00240557.
PR 29-JAN-1999; 99US-00241101.
PR 11-FEB-1999; 99US-00248673.
PR 11-FEB-1999; 99US-00248674.
PR 02-MAR-1999; 99US-0122450P.
PR 02-MAR-1999; 99US-0122452P.
PR 02-MAR-1999; 99US-0122560P.
PR 02-MAR-1999; 99US-0122565P.
PR 02-MAR-1999; 99US-0122566P.
PR 11-MAR-1999; 99US-00267719.
PR 28-JAN-2000; 2000US-00494096.
PR 16-MAR-2001; 2001US-0276822P.
XX
(CAPL/) CAPLAN M J.
PA (SOSI/) SOSIN H B.
PA (SAMP/) SAMPSON H.
PA (BANN/) BANNON G A.
PA (BURK/) BURKS A W.
PA (COCK/) COCKRELL G.
PA (COMP/) COMPADRE C M.
PA (CONN/) CONNAUGHTON C.
PA (HELM/) HELM R M.
PA (KING/) KING N E.
PA (KOPP/) KOPPER R A.
PA (MALE/) MALEKI S J.

PA (RABJ/) RABJOHN P A.
PA (SHIN/) SHIN D S.
XX (STAN/) STANLEY J S.
XX
PI Caplan MJ, Sosin HB, Sampson H, Bannan GA, Burks AW, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI Rabjohn PA, Shin DS, Stanley JS;
XX
DR WPI: 2003-875632/81.
DR N-PSDB; ADG27517.
XX
PT New modified anaphylactic food allergen comprising a cysteine residue
PT which has been modified so that it cannot participate in the disulfide
PT bond, useful for treating allergic reactions or wounds.
XX
PS Claim 27; SEQ ID NO 63; 194pp; English.
XX
CC The invention relates to a modified anaphylactic food allergen whose
CC amino acid sequence is substantially identical to that of a natural
CC anaphylactic food allergen. The natural anaphylactic food allergen
CC includes at least one cysteine residue that participates in a disulphide
CC bond when the natural anaphylactic food allergen is in its native
CC conformation, except that the cysteine residue has been modified so that
CC it cannot participate in the disulphide bond. Also included are a method
CC of making a modified anaphylactic food allergen, a nucleotide molecule
CC encoding a modified anaphylactic food allergen defined above, a
CC nucleotide molecule for causing a site specific mutation in a gene
CC encoding a natural anaphylactic food allergen, a transgenic plant or
CC animal expressing a modified anaphylactic food allergen defined above, a
CC method of treating an individual by reducing the clinical response to a
CC natural anaphylactic food allergen by administering a modified
CC anaphylactic food allergen and an isolated fragment of peanut allergen
CC Ara h 1, comprising at least 10 consecutive amino acids of ADG27464 or
CC ADG27465. About 10-17% of the amino acids have been modified in at least
CC one IGE epitope or all the IGE epitopes recognised when the natural
CC anaphylactic food allergen is contacted with serum IGE from individual(s)
CC allergic to the natural anaphylactic food allergen. The invention
CC discloses Peanut allergens Ara h1, Ara h2, Ara h3 (and their encoding
CC cDNAs), Soybean Glycinin A2B1a and IGE-binding epitopes of the English
CC walnut allergen Jug n1. The modified anaphylactic food allergen can be
CC used for treating allergic reactions or wounds. The present sequence
CC represents a Peanut allergen of the invention (or its fragment).
XX
SQ Sequence 157 AA;
Alignment Scores:
Pred. No.: 9.32e-90 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 65.7% Indels: 0
DB: 7 Gaps: 0
US-10-728-323-2 (1-717) x ADG27518 (1-157)
QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCTCGTCCCGCCATCTCGAGGCAG 61
Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisSerIlaArgGln 20
QY 62 CAGTGGGAACCTCCAAGGAGACAGAGATGCCAGAGCCAGCTCGAGGGCGCACTGAGG 121
Db 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCCTCGGAGCAACATCTCATCGAGAGATCCACCTGACGAGGATTCATATGAACGGAC 181
Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCGTGA 241
Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
QY 242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAACACCAAGG 301

XX New isolated nucleic acid encoding peanut allergen Ara h 1, useful for
PT obtaining diagnostics with optimal concentrations of each allergen, for
PT developing panels of mixtures of recombinant allergens, and in
PT immunotherapy.

XX Disclosure; Fig 27; 141pp; English.

XX This invention relates to a novel isolated nucleic acid molecule encoding
XX the peanut allergen Ara h I that has been identified in patients with
XX atopic dermatitis. Specifically, it refers to a monoclonal antibody
XX enzyme-linked immunosorbent assay (ELISA) used to detect and quantify the
XX peanut allergens, where the monoclonal antibodies have specificity for a
XX selected peanut allergen antigen. The present invention also describes
XX hybridomas that produce two-site monoclonal antibodies specific for a
XX peanut allergen that can be used in an ELISA to detect and determine the
XX concentration of a specific peanut allergen in a food product or in food
XX processing. Furthermore, it provides an expression vector and the
XX isolated nucleic acid or its fraction can include a diagnostic label.
XX Accordingly, these nucleic acids are useful for obtaining diagnostics
XX with optimal concentrations of each allergen or for developing panels of
XX mixtures of large numbers of recombinant allergens such that it can be
XX used in immunotherapy for the treatment of food hypersensitivity
XX reactions. This polypeptide sequence is the peanut Ara h II allergen
XX protein (P38 clone) of the invention.

XX Sequence 159 AA;

Alignment Scores:
Pred. No.: 9.37e-90 Length: 159
Score: 842.00 Matches: 157
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 65.7% Indels: 0
DB: 9 Gaps: 0

US-10-728-323-2 (1-717) x ADV97603 (1-159)

QY 2 CTCACCATCTAGTAGCTCGCCCTTTCTCTCGCTGCCACGCATCTCGAGGCAG 61
DB 3 LeuThrIleLeuValIleLeuAlaLeuPheLeuLeuAlaAlaHisIleSerAlaArgGln 22
QY 62 CAGTGGGAATCTCAAGGAGACAGAAGATCCAGAGCCAGCTCGAGAGCGCGAACCCTGAGG 121
DB 23 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 42
QY 122 CCTCGGAGCAATCTCATCGAGAGATCCACGTCGACGAGGATTCATATGAACGGAC 181
DB 43 ProCysGluGlnHisLeuMetGlnTyrIleGlnArgAspGluAspSerTyrGluArgAsp 62
QY 182 CGGTACAGCCCTAGTCAGGATCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 241
DB 63 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 82
QY 242 TCCTCTCAGCACCACAGAGAGGTGTTCGAATGAGCTGAACGAGTTTGAGAACCAACCAAGG 301
DB 83 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 102
QY 302 TGCATGTGGAGCATTCGACATGATCATCGAGAACGACGATGAGTGTGAGGGGAGG 361
DB 103 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 122
QY 362 CAACAGGAGCAACAGTTCAAGAGGAGGAGCTCAGGAACCTGCCTCAACAGTGGCGCTTAGG 421
DB 123 GlnGlnGlnGlnPheIleYsArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 142
QY 422 GCACCAAGCGTTTGGACCTTGGACGTGGAAGTGGCGGCGGACGACAGATAC 472
DB 143 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 159

RESULT 9

AAB82383

ID AAB82383 standard; protein; 207 AA.

XX AAB82383;
AC 11-SEP-2003 (revised)
DT 23-JUL-2001 (first entry)
XX Peanut allergen Ara h2 gene product.
XX Peanut; allergen; Ara h2; transgenic plant; allergy.
XX Arachis hypogaea.

XX Key Location/Qualifiers
FH Peptide 1..21
FT /label= Signal_peptide
FT Protein 22..207
FT /label= Mature_protein

XX WO200136621-A2.

XX 25-MAY-2001.

XX 20-NOV-2000; 2000WO-US031657.

XX 19-NOV-1999; 99US-0167255P.

XX (UYAL-) UNIV ALABAMA A & M.

XX Dodo HW, Arntzen CJ, Konan KN, Viquez OM;

XX WPI; 2001-355630/37.

XX N-PSDB; AAF90336, AAF90337.

XX Producing transgenic peanut plants that produce allergen-free seeds,
XX useful in non-allergenic foods, by antisense or sense co-suppression of
XX allergen-encoding genes.

XX Example 1; Fig 2; 72pp; English.

XX The present sequence is that of the peanut allergenic protein (AP)
XX encoded by the Ara h2 gene (see AAF90336). The invention relates to a
XX method for producing a peanut plant having reduced, or undetectable, AP
XX content in its seed. A peanut plant cell is transformed with a DNA
XX construct containing an antisense AP gene and/or sense AP gene, or their
XX fragments, regenerated to plants, and fertile transgenic plants that
XX produce seeds with reduced AP content are identified. The AP sense or
XX antisense gene may comprise at least a portion of the Ara h2 gene
XX sequence. The seeds are useful for preparation of allergen-free foods.
XX Recombinant AP may be produced and used to produce antibodies useful for
XX detecting AP in foods, and for treatment or prevention of peanut allergy.
XX (Updated on 11-SEP-2003 to standardise OS field)

XX Sequence 207 AA;

Alignment Scores:
Pred. No.: 2.71e-89 Length: 207
Score: 838.50 Matches: 162
Percent Similarity: 90.3% Conservative: 5
Best Local Similarity: 87.6% Mismatches: 17
Query Match: 65.5% Indels: 1
DB: 4 Gaps: 1

US-10-728-323-2 (1-717) x AAB82383 (1-207)

QY 2 CTCACCATCTAGTAGCTCGCCCTTTCTCTCGCTGCCACGCATCTCGAGGCAG 61
DB 4 LeuThrIleLeuValIleLeuAlaLeuPheLeuLeuAlaAlaHisIleSerAlaArgGln 23
QY 62 CAGTGGGAATCTCAAGGAGACAGAAGATCCAGAGCCAGCTCGAGAGCGCGAACCCTGAGG 121
DB 24 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
QY 122 CCTCGGAGCAACATCTCTCATGCAGAAGATCCAACGTCGACGAGGATTCATATGAACGGGAC 181

Db 44 ProCysGlnGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 63
QY 182 CCCTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGCGCTGGA 241
Db 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 83
QY 242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAACACCAAGG 301
Db 84 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 103
QY 302 TGCATGTGCGAGCGATTGCAACAGATCATGGAGAACAGAGCGATAGTTGCGAGGGAGG 361
Db 104 CysMetCysGlnAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123
QY 362 CAACAGAGCAACAGATTCAGAGGGAGCTCAGGAACTTGCTCAACAGTGGCGCTTAGG 421
Db 124 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
QY 422 GCACCCAGCGTTGCGACTTGGACGTCGAAAGTGGCGGCGAGACAGATCTAAACACCT 481
Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyArgArgProArgIleProPro 163
QY 482 ATCTCAAAAGAAAGAAAGAAAGAAAGAAATAGCTTATATATAAGC---TATTATCTA 538
Db 164 IleLeuThrGlySerArgSerArgArgHisGlnSerProTyrGlyAsnArgArgTyrSer 183
QY 539 TGGTTATGTTTAGTT 553
Db 184 AlaMetCysLeuLeu 188
RESULT 10
AAV40973
ID AAV40973 standard; protein; 156 AA.
XX AC AAV40973;
XX DT 17-OCT-2003 (revised)
DT 06-DEC-1999 (first entry)
XX Ara h 2 protein fragment.
XX Peanut; allergen; Ara h 1; IgE; immunoglobulin E; epitope; Ara h 3;
KW allergic reaction; Ara h 2.
XX Arachis hypogaea.
XX WO945961-A1.
XX 16-SEP-1999.
XX 12-MAR-1999; 99WO-US005494.
XX 12-MAR-1998; 98US-0077763P.
PR 11-MAR-1999; 99US-0007763.
XX (UYAR-) UNIV ARKANSAS.
XX Burks W, Helm RM, Cockrell G, Bannon GA, Stanley JS, Shin DS;
PI Sampson H, Compadre CM, Huang SK, Maleki SJ, Kopper RA;
XX WPI; 1999-551218/46.
XX Tertiary structure of peanut allergen Ara h 1 for protection of a host
PT animal from allergic reaction.
XX Disclosure; Page 104; 193pp; English.
XX The invention provides a tertiary structure for the peanut allergen Ara H
CC 1. The Ara H 1 allergen is found to contain 23 linear IgE-binding
CC epitopes. The invention also provides an isolated recombinant peanut
CC allergen designated Ara h 3 and a nucleotide molecule encoding the peanut
CC allergen Ara h 3. Molecules of the invention are used to protect a host

CC animal from allergic reaction, particularly using a modified allergen
CC which is less reactive with IgE. The invention may also be used to ensure
CC that the allergen is not introduced into genetically modified food. The
CC present sequence represents an Ara h 2 protein fragment. (Updated on 17-
CC OCT-2003 to standardise OS field)
XX Sequence 156 AA;
SQ Alignment Scores: 2.42e-88 Length: 156
Pred. No.: 830.00 Matches: 155
Score: 830.00
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 64.8% Indels: 0
DB: Gaps: 0
US-10-728-323-2 (1-717) x AAV40973 (1-156)
QY 2 CTCACCATAGTAGCCCTCGCCCTTTCCTCTCGCTGCCACGCATCTGCGAGGCGAG 61
Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAACTCCAAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGGCGAAACCTGAGG 121
Db 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCTCGCAGCAACATCTCATGCAAGATCCAAACGTGACGAGGATTCATATGACGGGAC 181
Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGGCGCTGGA 241
Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
QY 242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAACACCAAGG 301
Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
QY 302 TGCATGTGCGAGCGATTGCAACAGATCATGGAGAACAGAGCGATAGTTGCGAGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGAGCAACAGATTCAGAGGGAGCTCAGGAACTTGCTCAACAGTGGCGCGAGAGAC 466
Db 121 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACCCAGCGTTGCGACTTGGACGTCGAAAGTGGCGGCGAGAGAC 466
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAsp 155
RESULT 11
ADV97553
ID ADV97553 standard; protein; 157 AA.
XX AC ADV97553;
XX DT 10-MAR-2005 (first entry)
XX Peanut Ara h II allergen (clone P38) deduced protein sequence.
DE DE
XX allergen; atopic dermatitis; diagnostic; immunotherapy;
KW immediate type hypersensitivity; Ara h II; dermatological; antiallergic;
KW immunosuppressive; antibody engineering; seed storage protein.
XX Arachis hypogaea.
XX US6835824-B1.
XX 28-DEC-2004.
XX 13-NOV-1998; 98US-00191593.
XX 29-DEC-1995; 95US-0009455P.
PR

PR 23-SEP-1996; 96US-00717933.
XX (UYAR-) UNIV ARKANSAS.
PI Burks AW, Stanley JS, Bannon GA, Cockrell G, Helm RM;
XX WPI; 2005-045982/05.
XX
XX New isolated nucleic acid encoding peanut allergen Ara h 1, useful for
PT obtaining diagnostics with optimal concentrations of each allergen, for
PT developing panels of mixtures of recombinant allergens, and in
PT immunotherapy.
XX
XX Disclosure; SEQ ID NO 21; 141pp; English.
XX
XX This invention relates to a novel isolated nucleic acid molecule encoding
CC the peanut allergen Ara h 1 that has been identified in patients with
CC atopic dermatitis. Specifically, it refers to a monoclonal antibody
CC enzyme-linked immunosorbent assay (ELISA) used to detect and quantify the
CC peanut allergens, where the monoclonal antibodies have specificity for a
CC selected peanut allergen antigen. The present invention also describes
CC hybridomas that produce two-site monoclonal antibodies specific for a
CC peanut allergen that can be used in an ELISA to detect and determine the
CC concentration of a specific peanut allergen in a food product or in food
CC processing. Furthermore, it provides an expression vector and the
CC isolated nucleic acid or its fraction can include a diagnostic label.
CC Accordingly, these nucleic acids are useful for obtaining diagnostics
CC with optimal concentrations of each allergen or for developing panels of
CC mixtures of large numbers of recombinant allergens such that it can be
CC used in immunotherapy for the treatment of food hypersensitivity
CC reactions. This polypeptide sequence is the deduced peanut Ara h II
CC allergen protein (P38 clone) of the invention.
XX
XX Sequence 157 AA;

Alignment Scores:
Pred. No.: 5.48e-88 Length: 157
Score: 827.00 Matches: 155
Percent Similarity: 98.7% Conservative: 0
Best Local Similarity: 98.7% Mismatches: 2
Query Match: 64.6% Indels: 0
DB: 9 Gaps: 0

US-10-728-323-2 (1-717) x ADV97553 (1-157)

QY 2 CTCACCATAGTACCCCTCGCCCTTTCTCTCGTCCGACGATCGAGGCGAG 61
Db 1 LeuThrIleuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAATCCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGCGAACCTTGAGG 121
Db 21 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCTGGGAGCAATCTCATGCAGAGATCCAACTGAGAGGATTCATATGAACGGGAC 181
Db 41 ProCysGluGlnHisIleuMetGlnIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CCGTACAGCCCTAGTCAGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 241
Db 61 ProSerTyrProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
QY 242 TCCCTCTCAGCACCAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAACCAACCAAGG 301
Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
QY 302 TGCATGTGCGAGCATTCGACAGATCATGGAGAACCCAGAGCGATAGTTTGCAGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGGACCAAGTTCAAGAGGAGCTCAGGAATTCCTCCACAGTGGCGGCTTAGG 421
Db 121 GlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140

QY 422 GCACACACAGCGTTGCGACTTGACCTCGAAAGCTGCGCGCAGACAGATAC 472
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157
RESULT 12
ADO38314
ID ADO38314 standard; peptide; 156 AA.
XX
XX ADO38314;
AC
XX 15-JUL-2004 (first entry)
DT
XX
XX Ara h 2 MHC-class II-presented epitope #1.
DE
XX
XX Antibacterial; Virucide; Fungicide; Antiparasitic; Antiarthritic;
KW Antirheumatic; Neuroprotective; Antiinflammatory; Dermatological;
KW Immunosuppressive; Anti-diabetic; Antithyroid; Antiasthmatic;
KW Antiallergic; Cystostatic; Antipsoriatic; Gene therapy; Vaccine;
KW MHC Class II; Ii-key motif; immune response; anthrax; EBOLA; HIV;
KW influenza; vaccinia virus; infection; bacterium; virus; parasite; fungus;
KW rickettsia; rheumatoid arthritis; multiple sclerosis;
KW lupus erythematosus; diabetes mellitus; myasthenia gravis;
KW autoimmune thyroiditis; scleroderma; dermatomyositis; pemphigus; asthma;
KW allergic rhinitis; topical dermatitis; colitis; cancer; psoriasis;
KW adenoma; peanut; Ara h 2.
XX
XX Arachis hypogaea.
OS
XX US2004058881-A1.
PN
XX 25-MAR-2004.
PD
XX
XX 24-SEP-2002; 2002US-00253286.
PF
XX
XX 24-SEP-2002; 2002US-00253286.
PR
XX
XX (ANTI-) ANTIGEN EXPRESS INC.
PA
XX Humphreys RE, Xu M;
PI
XX WPI; 2004-294259/27.
DR
XX
XX New non-naturally occurring protein or polypeptide modified by
PT recombinant DNA techniques, useful for treating multiple sclerosis,
PT diabetes mellitus, myasthenia gravis, scleroderma, allergic rhinitis,
PT colitis, cancer or psoriasis.
XX
XX Example 2; Page 18; 90pp; English.

XX The invention relates to a non-naturally occurring protein or polypeptide
CC (i) modified by recombinant DNA techniques comprising: a C-terminal
CC element comprising an MHC Class II-presented epitope; an N-terminal
CC element comprising an Ii-key motif; and an intervening element comprising
CC a sequence of 4-11 amino acid residues where the modification by
CC recombinant DNA techniques taking place within elements (b) and (c). Also
CC described are methods for: suppressing or enhancing an immune response
CC directed toward an MHC (major histocompatibility complex) Class II-
CC presented epitope of interest. Suppressing an immune response directed
CC toward an MHC Class II-presented epitope of interest comprises: providing
CC a nucleic acid sequence encoding the MHC Class II-presented epitope of
CC interest, the nucleic acid sequence encoding an Ii-key motif located 4-11
CC amino acids upstream from the N-terminal residue of the MHC Class II-
CC presented epitope of interest; and modifying the Ii-key motif to decrease
CC its conformance to the archetypal Ii-key regulatory motif. Enhancing an
CC immune response directed toward an MHC Class II-presented epitope of
CC interest comprises: providing a nucleic acid sequence encoding the MHC
CC Class II-presented epitope of interest, the nucleic acid sequence lacking
CC an Ii-key motif located 4-11 amino acids upstream from the N-terminal
CC residue of the MHC Class II-presented epitope of interest; and modifying
CC the nucleic acid sequence to introduce an Ii-key motif appropriately
CC spaced from the MHC Class II-presented epitope. The protein or
CC polypeptide of interest corresponds to a protein or polypeptide encoded
CC by an infectious pathogen selected from anthrax, EBOLA, HIV or influenza,

CC preferably vaccinia virus. The non-naturally occurring protein or
 CC polypeptide (1) modified by recombinant DNA techniques is useful for
 CC treating infectious diseases caused or associated with infection by a
 CC bacterium, virus, parasite, fungus, rickettsia or other infectious
 CC agents. It is also useful for treating rheumatoid arthritis, multiple
 CC sclerosis, lupus erythematosus, diabetes mellitus, myasthenia gravis,
 CC autoimmune thyroiditis, scleroderma, dermatomyositis, pemphigus, asthma,
 CC allergic rhinitis, topical dermatitis, colitis, cancer, psoriasis or
 CC adenomas. The present sequence represents the amino acid sequence of a
 CC peanut allergen Ara h 2 MHC class II-presented epitope used in the
 CC invention.

XX Sequence 156 AA;

Alignment Scores:
 Pred. No.: 4.8e-87 Length: 156
 Score: 819.00 Matches: 153
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 63.9% Indels: 0
 DB: 8 Gaps: 0

US-10-728-323-2 (1-717) x ADO38314 (1-156)

Qy 2 CTCACCATAGTACGCTCGCCCTTTCTCTCGTCCCGCCACGCATCTGGAGGCGAG 61
 Db 4 LeuThrIleuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaAraGln 23
 Qy 62 CAGTGGGAATCCAAAGGAGACAGAAAGATGCAGAGCCAGCTCGAGAGGGCGAACCTCAGG 121
 Db 24 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
 Qy 122 CCTCGGAGCAACATCTCATGCAAGATCCAACTGACGAGATTCATATGAAACGGGAC 181
 Db 44 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgasp 63
 Qy 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGCGCTTGA 241
 Db 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 83
 Qy 242 TCCTCTCAGACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTGAGAACCAACCAAGG 301
 Db 84 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 103
 Qy 302 TGCATGTGCGAGGATTCGCAACAGATCATGAGAGAACGAGAGCGATAGGTTGCGGGAGG 361
 Db 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123
 Qy 362 CAACAGGAGCAACAGTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTCGGCCTTAGG 421
 Db 124 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
 Qy 422 GCACCAAGCGTTGCGACTTGAGCGTTCGAAAGTGGCGCG 460
 Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyGly 156

RESULT 13

ID ABU52576 standard; protein; 157 AA.

AC ABU52576;

XX 10-MAR-2003 (first entry)

DT Peanut Ara h2 mutant 5 (Q20A/Q31A/D60A/D67A).

DE Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;
 KW munein; anaphylactic food allergen; antiallergenic; vaccine;
 KW wound healing.

XX Homo sapiens.

XX WO200274250-A2.

PN

XX 26-SEP-2002.

PD 18-MAR-2002; 2002WO-US009108.

XX 16-MAR-2001; 2001US-0276822P.

PR 18-MAR-2002; 2002US-00276822.

XX (PANA-) PANACEA PHARM.

XX Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
 PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
 PI Rabinjohn PA, Shin DS, Stanley JS;

XX WPI; 2003-018765/01.

XX New modified anaphylactic food allergen, useful for preventing or
 PT treating allergic reactions associated with e.g. anaphylactic allergens.

XX Example 13; Page: 300pp; English.

XX The invention relates to a modified anaphylactic food allergen has an
 CC amino acid sequence that is substantially identical to that of natural
 CC anaphylactic food allergen, except for a cysteine residue that has been
 CC modified so that it cannot participate in the disulphide bond. The
 CC modification may also comprise mutation of the IgE binding sites to
 CC reduce allergenicity. Also included are: (1) a method of making a
 CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
 CC or for causing a site specific mutation in the modified anaphylactic food
 CC allergen; (3) a transgenic plant or animal expressing the modified
 CC anaphylactic food allergen; (4) a method of treating an individual by
 CC reducing the clinical response to a natural anaphylactic food allergen;
 CC and an isolated fragment of peanut allergen Ara h 1. The modified
 CC anaphylactic food allergen is useful for preventing or treating allergic
 CC reactions associated with any natural allergen such as food, insect,
 CC rubber or preferably anaphylactic allergens. It is also useful for
 CC treating wounds in mammals such as bovine, canine, feline, caprine,
 CC ovine, porcine, murine or equine species. The present sequence is a food
 CC allergen, mutated to alter its IgE binding characteristics. Note: The
 CC present sequence is not shown in the specification but was created by the
 CC indexer using information provided in the specification

XX Sequence 157 AA;

Alignment Scores:

Pred. No.: 1.87e-86 Length: 157
 Score: 814.00 Matches: 153
 Percent Similarity: 97.5% Conservative: 0
 Best Local Similarity: 97.5% Mismatches: 4
 Query Match: 63.5% Indels: 0
 DB: 6 Gaps: 0

US-10-728-323-2 (1-717) x ABU52576 (1-157)

Qy 2 CTCACCATAGTACGCTCGCCCTTTCTCTCGTCCCGCCACGCATCTGGAGGCGAG 61
 Db 1 LeuThrIleuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaAraGala 20
 Qy 62 CAGTGGGAATCCAAAGGAGACAGAAAGATGCAGAGCCAGCTCGAGAGGGCGAACCTCAGG 121
 Db 21 GlnTrpGluLeuGlnGlyAspArgArgCysAlaSerGlnLeuGluArgAlaAsnLeuArg 40
 Qy 122 CCTCGGAGCAACATCTCATGCAAGATCCAACTGACGAGATTCATATGAAACGGGAC 181
 Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgala 60
 Qy 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTTGA 241
 Db 61 ProTyrSerProSerGlnAlaProTyrSerProSerProTyrAspArgGlyAlaGly 80
 Qy 242 TCCTCTCAGACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTGAGAACCAACCAAGG 301
 Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 100

QY 302 TGCATGTGCGAGGATTGCAACAGATCATGGAGAACAGAGCGGATAGTTGCGAGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGAGCAACAGTTCAAGAGGAGGCTCAGGAACCTGCTCAACAGTGCAGGCTTAGG 421
Db 121 GlnGlnGlnGlnGlnPheGlyArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACCAACAGCGTTGCCACTTGGACGTCGAAGTGGCGGCGAGACAGATAC 472
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157
RESULT 14
ADM12096
ID ADM12096 standard; protein; 156 AA.
XX AC ADM12096;
XX DT 20-MAY-2004 (first entry)
XX Arachis hypogaea 2 (Ara h2) protein.
DE Arachis hypogaea 2 (Ara h2) protein.
XX antigen presentation enhancing hybrid polypeptide; mammalian II-Key;
KW MHC class II; antibacterial; virucide; fungicide; antirheumatic;
KW antiarthritic; neuroprotective; dermatological; immunosuppressive;
KW antiinflammatory; antidiabetic; antithyroid; immune;
KW rheumatoid arthritis; multiple sclerosis; lupus erythematosus;
KW diabetes mellitus; myasthenia gravis; autoimmune thyroiditis;
KW scleroderma; dermatomyositis; pemphigus.
XX Arachis hypogaea.
XX US2003235594-A1.
XX 25-DEC-2003.
XX 17-SEP-2002; 2002US-00245871.
PR 14-SEP-1999; 99US-00396813.
PR 17-JUL-2002; 2002US-00197000.
XX (ANTI-) ANTIGEN EXPRESS INC.
XX Humphreys R, Xu M;
XX WPI; 2004-070554/07.
XX Novel II-Key/antigen presentation enhancing hybrid polypeptide, useful
PT for treating infections, rheumatoid arthritis, multiple sclerosis, lupus
PT erythematosus and diabetes mellitus.
XX Example 2; Page 19; 87pp; English.
XX The invention relates to a novel antigen presentation enhancing hybrid
CC polypeptide. The novel polypeptide has an N-terminal element consisting
CC of 4-16 residues of a mammalian II-Key peptide and its non-N-terminal
CC deletion modifications, a chemical structure covalently linking the N-
CC terminal element to an MHC class II-presented epitope of a C-terminal
CC element. The C-terminal element comprises an antigenic epitope, which
CC binds to an antigenic peptide binding site of an MHC class II molecule.
CC The antigen presentation enhancing hybrid polypeptide has the following
CC activities: antibacterial, virucide, fungicide, antirheumatic,
CC antiarthritic, neuroprotective, dermatological, immunosuppressive,
CC antiinflammatory, antidiabetic, and antithyroid. The antigen presentation
CC enhancing hybrid polypeptide is useful for modulating the immune response
CC in an individual and for treating infections (such as bacteria, virus,
CC parasite and fungus), rheumatoid arthritis, multiple sclerosis, lupus
CC erythematosus, diabetes mellitus, myasthenia gravis, autoimmune
CC thyroiditis, scleroderma, dermatomyositis and pemphigus. This sequence
CC represents a mammalian II key related protein of the invention.
XX Sequence 156 AA;
SQ

Alignment Scores:
Pred. No.: 3 21e-86 Length: 156
Score: 812.00 Matches: 152
Percent Similarity: 99.3% Conservatives: 0
Best Local Similarity: 99.3% Mismatches: 1
Query Match: 63.4% Indels: 0
DB: 8 Gaps: 0
US-10-728-323-2 (1-1717) x ADM12096 (1-156)
QY 2 CTCACATACATAGTAGCCCTCCCTTTCTCTCTCGTCCACGACATCTCGAGGAGG 61
Db 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 23
QY 62 CAGTGGGAACCTCAAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGCGAACCCTGAGG 121
Db 24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
QY 122 CCCTCGAGCAACATCTCATGCAGAAAGATCCCAACGTGACGAGGATTTCATATGAACGGGAC 181
Db 44 ProCysGluGlyHisLeuMetGlnIleGlnArgAspGluAspSerTyrGluArgAsp 63
QY 182 CCGTACAGCCCTAGTCAGATCCGTACAGCCCTAGTCCATATGATCGAGAGGCGCTGGA 241
Db 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 83
QY 242 TCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGACAGCTTTGAGAACCAACCAAGG 301
Db 84 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnGlnArg 103
QY 302 TGCATGTGCGAGGATTGCAACAGATCATGGAGAACAGAGCGATAGTTGCGAGGAGG 361
Db 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123
QY 362 CAACAGGAGCAACAGCTTCAAGAGGAGCTCAGGAACCTTGCCTCAACAGTGCAGGCTTAGG 421
Db 124 GlnGlnGluGlnGlnPheGlyArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
QY 422 GCACCAACAGCGTTGCGACTTGGACGTCGAAGTGGCGGCGAGACAGATAC 460
Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyGly 156
RESULT 15
ABU52575
ID ABU52575 standard; protein; 157 AA.
XX AC ABU52575;
XX DT 10-MAR-2003 (first entry)
XX Peanut Ara h2 mutant 4 (E35A/P41A/D60A/D67A).
XX Allergy; Ara h1; Ara h2; Ara h3; IgE binding site; peanut; mutant;
KW mutein; anaphylactic food allergen; antiallergenic; vaccine;
XX wound healing.
XX Homo sapiens.
XX OS
XX WO200274250-A2.
XX 26-SEP-2002.
XX 18-MAR-2002; 2002WO-US0009108.
XX 16-MAR-2001; 2001US-0276822P.
XX 18-MAR-2002; 2002US-00276822.
XX (PANA-) PANACEA PHARM.
XX Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;
PI Compadre CM, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;
PI RabJohn PA, Shin DS, Stanley JG;

Search completed: May 15, 2006, 21:44:34
Job time : 110.391 secs

XX WPI; 2003-018765/01.
XX
XX New modified anaphylactic food allergen, useful for preventing or
PT treating allergic reactions associated with e.g. anaphylactic allergens.
XX
XX Example 13; Page; 300pp; English.
PS
XX
XX The invention relates to a modified anaphylactic food allergen has an
CC amino acid sequence that is substantially identical to that of natural
CC anaphylactic food allergen, except for a cysteine residue that has been
CC modified so that it cannot participate in the disulphide bond. The
CC modification may also comprise mutation of the IgE binding sites to
CC reduce allergenicity. Also included are: (1) a method of making a
CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
CC or for causing a site specific mutation in the modified anaphylactic food
CC allergen; (3) a transgenic plant or animal expressing the modified
CC anaphylactic food allergen; (4) a method of treating an individual by
CC reducing the clinical response to a natural anaphylactic food allergen;
CC and an isolated fragment of peanut allergen Ara h 1. The modified
CC anaphylactic food allergen is useful for preventing or treating allergic
CC reactions associated with any natural allergen such as food, insect,
CC rubber or preferably anaphylactic allergens. It is also useful for
CC treating wounds in mammals such as bovine, canine, feline, caprine,
CC ovine, porcine, murine or equine species. The present sequence is a food
CC allergen, mutated to alter its IgE binding characteristics. Note: The
CC present sequence is not shown in the specification but was created by the
CC indexer using information provided in the specification
XX
XX Sequence 157 AA;

Alignment Scores:
Pred. No.: 3.22e-86 Length: 157
Score: 812.00 Matches: 153
Percent Similarity: 97.5% Conservative: 0
Best Local Similarity: 97.5% Mismatches: 4
Query Match: 63.4% Indels: 0
DB: 6 Gaps: 0

US-10-728-323-2 (1-717) x ABUS2575 (1-157)

QY	2	CTCACCATACTAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCATCTCGAGGCGAG	61
DB	1	LeuThrIleuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln	20
QY	62	CAGTGGGAACCTCCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGGGCGAACCCTGAGG	121
DB	21	GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuAlaArgAlaAsnLeuArg	40
QY	122	CCCTGCGAGCAACATCTCATGCGAGAGATCCACGTGACGAGGATTCATATGACGGGAC	181
DB	41	AlaCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAla	60
QY	182	CCGTACAGCCCTAGTACAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCTGGA	241
DB	61	ProTyrSerProSerGlnAlaProTyrSerProTyrAspArgArgGlyAlaGly	80
QY	242	TCCTCTCAGACCCAGAGAGGTGTTCGAATGAGTGAACGAGTTTGAGAACAAACCAAGG	301
DB	81	SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg	100
QY	302	TGCATGTCCGAGGATTCACACAGATCATGGAGNACCAGAGCGATAGTTTCAGGGGAGG	361
DB	101	CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg	120
QY	362	CAACAGGAGCAACAGTTTCAAGAGGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGG	421
DB	121	GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg	140
QY	422	GCACCACAGCGTTTCGACCTTGACGTCGAAAGTGGCGGCGAGACAGATAC	472
DB	141	AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr	157

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: May 15, 2006, 22:18:32 ; Search time 4.3124 Seconds
(without alignments)
4123.811 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 1281
Sequence: 1 gctcaccatactagtagcccc.....taaaagatcatgttttgtt 717

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delopt 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q=/abs/ABSWEB_spool/US10728323/runat_15052006_172137_22463/app_query.fasta_1
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=n2p.ra -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs02p -USER=US10728323 @CEN 1 1 131 @runat 15052006 172137 22463
-NCPU=6 -ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/iaa/5 COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PTUS COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	842	65.7	157	2	US-09-106-872A-22
2	842	65.7	157	2	US-09-191-593-10
3	842	65.7	159	2	US-09-106-872A-2
4	838.5	65.5	207	2	US-09-715-036-2
5	827	64.6	157	2	US-09-191-593-21
6	243	19.0	158	1	US-08-618-911-4
7	243	19.0	158	2	US-08-618-911-2
8	243	19.0	158	2	US-09-531-727-2
9	241.5	18.9	155	1	US-08-618-911-2
10	229	17.9	158	1	US-08-618-911-6
11	156	12.2	28	2	US-09-191-593-53
12	156	12.2	158	1	US-08-670-186-4

13	151	11.8	158	1	US-08-670-186-6	Sequence 6, Appli
14	139	10.9	155	1	US-08-670-186-2	Sequence 2, Appli
15	139	10.9	184	1	US-08-453-924-3	Sequence 3, Appli
16	138	10.8	28	2	US-09-191-593-58	Sequence 58, Appl
17	136.5	10.7	164	2	US-09-627-536-13	Sequence 13, Appl
18	129.5	10.1	150	2	US-10-442-174A-1	Sequence 1, Appli
19	118	9.2	30	2	US-09-106-872A-16	Sequence 16, Appl
20	118	9.2	30	2	US-09-191-593-57	Sequence 57, Appl
21	106.5	8.3	174	2	US-09-645-593-7	Sequence 7, Appli
22	106.5	8.3	1088	2	US-09-233-857-13	Sequence 13, Appl
23	105.5	8.2	1099	2	US-09-442-100-2	Sequence 2, Appli
24	105.5	8.2	1099	2	US-08-939-108-2	Sequence 2, Appli
25	105.5	8.2	1099	2	US-09-442-103-2	Sequence 2, Appli
26	102	8.0	316	2	US-08-098-327E-31	Sequence 31, Appl
27	102	8.0	316	2	US-08-463-625-31	Sequence 31, Appl
28	100	7.8	17	2	US-09-191-593-55	Sequence 55, Appl
29	100	7.8	17	2	US-09-191-593-61	Sequence 61, Appl
30	100	7.8	17	2	US-09-191-593-62	Sequence 62, Appl
31	100	7.8	224	2	US-09-270-767-44901	Sequence 44901, A
32	99.5	7.8	148	2	US-09-083-852A-4	Sequence 4, Appli
33	99.5	7.8	148	2	US-09-083-852A-6	Sequence 6, Appli
34	99.5	7.8	148	2	US-09-489-674B-6	Sequence 6, Appli
35	97	7.6	185	2	US-09-252-991A-23490	Sequence 23490, A
36	95	7.4	32	1	US-08-158-704-1	Sequence 1, Appli
37	95	7.4	32	1	US-08-610-424B-1	Sequence 1, Appli
38	95	7.4	32	2	US-09-336-463-1	Sequence 1, Appli
39	95	7.4	32	2	US-09-191-593-1	Sequence 1, Appli
40	95	7.4	152	2	US-09-252-991A-19284	Sequence 19284, A
41	94	7.3	43	2	US-09-303-814-1	Sequence 1, Appli
42	94	7.3	500	2	US-09-265-630-13	Sequence 13, Appl
43	93.5	7.3	415	2	US-09-252-991A-18252	Sequence 18252, A
44	93	7.3	537	2	US-09-252-991A-27024	Sequence 27024, A
45	92.5	7.2	788	1	US-08-918-914-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-106-872A-22
; Sequence 22, Application US/09106872A
; Patent No. 6486311
; GENERAL INFORMATION:
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina E.
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricki M.
; APPLICANT: Bannan, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106,872A
; CURRENT FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: PCT/US96/15222
; PRIOR FILING DATE: 1996-09-23
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-09-106-872A-22

Alignment Scores:
Pred. No.: 2,528-86 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 65.7% Indels: 0
DB: 2 Gaps: 0

US-10-728-323-2 (1-717) x US-09-106-872A-22 (1-157)

Db 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCTGGAGCAACATCTCATGACAGAGATCCAACTGACGAGGATTTCATATCAACGGGAC 181
Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACACCCCTAGTCCATATGATCGGAGCGCTGGA 241
Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
QY 242 TCCTCTCAGCACACAGAGAGGTTCGCAATGAGCTGAACAGCTTGAGAACACCAAGG 301
Db 81 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 100
QY 302 TGCATGTGCGAGGATTCGACAGATCATCGAGAACCAAGAGCGATAGGTTGACGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGGACCAACAGTTCAGAGGGAGCTCAGAACTTGCCCTCAACAGTCCGCTTAGG 421
Db 121 GlnGlnGluGlnGlnPheLysArgGluLeuA:GAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACACAGCGTTGCGACTTGACGCTCGAAAGTGGCGGCGAGACAGATAC 472
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157
RESULT 3
US-09-106-872A-2
; Sequence 2, Application US/09106872A
; Patent No. 6486311
; GENERAL INFORMATION:
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina E.
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricki M.
; APPLICANT: Bannon, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106,872A
; CURRENT FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: PCT/US96/15222
; PRIOR FILING DATE: 1996-09-23
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Arachis hypogaea
; FEATURE:
; OTHER INFORMATION: Amino Acids 17-26 are Ara H 2 Binding Epitopes,
; OTHER INFORMATION: Peptide 1
; OTHER INFORMATION: Amino Acids 23-32 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 2
; OTHER INFORMATION: Amino Acids 29-38 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 3
; OTHER INFORMATION: Amino Acids 41-50 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 4
; OTHER INFORMATION: Amino Acids 51-60 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 5
; OTHER INFORMATION: Amino Acids 59-68 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 6
; OTHER INFORMATION: Amino Acids 67-76 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 7
; OTHER INFORMATION: Amino Acids 117-126 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 8
; OTHER INFORMATION: Amino Acids 129-138 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 9
; OTHER INFORMATION: Amino Acids 145-154 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 10
; OTHER INFORMATION: Amino Acids 13-22 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 1, Table 12

; OTHER INFORMATION: Amino Acids 27-36 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 2, Table 12
; OTHER INFORMATION: Amino Acids 59-68 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 3, Table 12
; OTHER INFORMATION: Amino Acids 81-90 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 4, Table 12
; OTHER INFORMATION: Amino Acids 91-100 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 5, Table 12
; OTHER INFORMATION: Amino Acids 145-154 are Ara H 2 binding epitopes,
; OTHER INFORMATION: peptide 6, Table 12
US-09-106-872A-2
Alignment Scores:
Pred. No.: 2,53e-86 Length: 159
Score: 842.00 Matches: 157
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 65.7% Indels: 0
DB: 2 Gaps: 0
US-10-728-323-2 (1-717) x US-09-106-872A-2 (1-159)
QY 2 CTCACCATACTAGTAGCCCTCGCCCTTTTCCTCTCGCTGCCACGCGCATCTCGGAGGAG 61
Db 3 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 22
QY 62 CAGTGGGAATCCCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAGG 121
Db 23 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 42
QY 122 CCCTCGAGCAACATCTCATGCGAAGATCCAACTGACGAGGATTCATATGAACGGGAC 181
Db 43 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 62
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGCGCTGGA 241
Db 63 ProTyrSerProSerGlnAspProTyrSerProTyrAspArgGlyAlaGly 82
QY 242 TCCTCTCAGCACACAGAGAGGTGTTCGAATGAGCTGAAACGAGTTCGAGAACCAACAGG 301
Db 83 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnGlnArg 102
QY 302 TGCATGTGCGAGGATTCGACAGATCATGAGAAACCAAGAGCGATAGGTTGACGGGAGG 361
Db 103 CysMetCysGluAlaLeuGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 122
QY 362 CAACAGGACCAACAGTTCAGAGGGAGCTCAGAACTTGCCCTCAACAGTCCGCTTAGG 421
Db 123 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 142
QY 422 GCACACAGCGTTGCGACTTGAGCTCGAAAGTGGCGGCGAGACAGATAC 472
Db 143 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 159
RESULT 4
US-09-715-036-2
; Sequence 2, Application US/09715036
; Patent No. 6943010
; GENERAL INFORMATION:
; APPLICANT: DODO, HORTENSE W.
; APPLICANT: ARNTZEN, CHARLES J.
; APPLICANT: KONAN, KOFFI N'DA
; APPLICANT: VIOUEZ, OLGA
; TITLE OF INVENTION: DOWN-REGULATION AND SILENCING OF ALLERGEN GENES IN
; FILE REFERENCE: 072121/0104
; CURRENT APPLICATION NUMBER: US/09/715,036
; CURRENT FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/167,255
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2

```
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-09-715-036-2

Alignment Scores:
Pred. No.: 6.95e-86 Length: 207
Score: 838.50 Matches: 162
Percent Similarity: 90.3% Conservative: 5
Best Local Similarity: 87.6% Mismatches: 17
Query Match: 65.5% Indels: 1
DB: 2 Gaps: 1

US-10-728-323-2 (1-717) x US-09-715-036-2 (1-207)
Qy 2 CTCACATAGTCTAGTCCCTCGCCCTTTCTCTCTGCTCCACGATCTCGAGGCGAG 61
Db 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 23
Qy 62 CAGTGGGAATCTCAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGGGGAACCTGAGG 121
Db 24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
Qy 122 CCCTGCGGCAACATCTCATGCAGAGATCCAAAGTGCAGGAGTTCATATGAACGGAC 181
Db 44 ProCysGlnGlnHisLeuMetGlnLysIleGlnArgAspGlnAspSerTyrgluArgasp 63
Qy 182 CCCTACAGCCCTAGTCCAGGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCTGGA 241
Db 64 ProTyrglnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 83
Qy 242 TCCTCTCAGCACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAACCAACGAAGG 301
Db 84 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 103
Qy 302 TGCAATGCGAGGATTCGACAGATCATGAGAACCCAGAGCGATAGTTGCGAGGGAGG 361
Db 104 CysMetCysGluAlaLeuGlnGlnIleMetGlnAsnGlnSerAspArgLeuGlnGlyArg 123
Qy 362 CAACAGAGCAACAGATTCAGAGAGGAGCTCAGAACTTGCTCAACAGTCGGCGCTTAGG 421
Db 124 GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnCysGlyLeuArg 143
Qy 422 GCACCAAGGCTTGGAGCTTGGAGCTGGAAGTGGCGGCGAGAGACAGATATAACACCT 481
Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgArgProArgIleProPro 163
Qy 482 ATCTCAAAAAAGAAAAAGAAAAAGAAATAGCTTATATATAAGC---TATTATCTA 538
Db 164 IleLeuThrGlySerArgSerArgArgHisGlnSerProTyrglyAsnArgArgTyrgSer 183
Qy 539 TGGTTATGTTTAGTT 553
Db 184 AlaMetCysLeuLeu 188

RESULT 5
US-09-191-593-21
; Sequence 21, Application US/09191593
; Patent No. 6835824
; GENERAL INFORMATION:
; APPLICANT: BURKS, A Wesley, HELM, Ricki M.
; APPLICANT: COCKRELL, Gael, STANLEY, J Steven,
; TITLE OF INVENTION: PEANUT ALLERGENS AND
; TITLE OF INVENTION: PEANUT ALLERGENS AND
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Head, Johnson & Kachigian
; STREET: 112 W. Center St., Suite 230
; CITY: Fayetteville
; STATE: Arkansas AR
; COUNTRY: United States of America
; ZIP: 72701

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS 6.2
; SOFTWARE: Wordperfect 6.0C
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/191,593
; FILING DATE: 13 NOVEMBER 1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/717,933
; FILING DATE: 23 SEPTEMBER 1996
; APPLICATION NUMBER: US 07/998,377
; FILING DATE: 30 DECEMBER 1992
; APPLICATION NUMBER: US 08/158,704
; FILING DATE: 29 NOVEMBER 1993
; APPLICATION NUMBER: US 60/009,455
; FILING DATE: 29 DECEMBER 1995
; APPLICATION NUMBER: US 08/610,424
; FILING DATE: 04 MARCH 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ALEXANDER, DANIEL R
; REGISTRATION NUMBER: 32,604
; REFERENCE/DOCKET NUMBER: ARK00895601B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (501) 582-9111
; TELEFAX: (501) 521-4931
; TELEX: No. 6835824 applicable
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 157 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. 6835824 applicable
; TOPOLOGY: unknown
; MOLECULE TYPE: glycoprotein
; DESCRIPTION: identified as Ara h 2 p38 deduced
; DESCRIPTION: sequence from nucleotide sequence reading frame.
; HYPOTHETICAL: No
; ANTI-SENSE: No. 6835824 applicable
; FRAGMENT TYPE: No. 6835824 applicable
; ORIGINAL SOURCE:
; ORGANISM: Arachis hypogaea
; STRAIN: Florunner
; INDIVIDUAL ISOLATE: Ara h II
; DEVELOPMENTAL STAGE:
; HAPLOTYPE: No. 6835824 applicable
; TISSUE TYPE:
; CELL TYPE: No. 6835824 applicable
; CELL LINE: No. 6835824 applicable
; ORGANELLE: No. 6835824 applicable
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: No. 6835824 applicable
; MAP POSITION: No. 6835824 applicable
; UNITS: No. 6835824 applicable
; FEATURE:
; NAME/KEY: final check
; LOCATION: 4..480
; IDENTIFICATION METHOD: By agreement with
; IDENTIFICATION METHOD: protein information and established
; IDENTIFICATION METHOD: consensus sequence
; OTHER INFORMATION: Seed storage protein and
; OTHER INFORMATION: allergen
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
```

```
;
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-09-191-593-21

Alignment Scores:
Pred. No.: 1,24e-84 Length: 157
Score: 827.00 Matches: 155
Percent Similarity: 98.7% Conservative: 0
Best Local Similarity: 98.7% Mismatches: 2
Query Match: 64.6% Indels: 0
DB: 2 Gaps: 0

US-10-728-323-2 (1-717) x US-09-191-593-21 (1-157)

QY 2 CTACATAGTACCTGCGCCCTTCTCTCTGCTGCCACGATCTCGAGGAG 61
Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAACTCCAGGAGACAGATGCCAGCCAGCTCGAGAGGGCGAACCTGAGG 121
Db 21 GlnTrpGluLeuGlnGlyArgArgGlyGlnSerGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCTCTGCGAGCAACATCTCATGCGAAGATCCAACTGACGAGGATTCATATGAACGGGAC 181
Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CCGTACAGCCCTAGTACAGATCCGTACAGCCCTAGTCCATATGATCGAGAGCGCTGGA 241
Db 61 ProSerTyrProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 80
QY 242 TCCTCTCAGCACCAAGAGGCTGTCGATGACGCTGACGAGCTTCGAGAACCAACGAGG 301
Db 81 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnGlnArg 100
QY 302 TGATGTCGAGGACATTCGACAGATCATGAGAACAGACGATAGGTTGACAGGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGAGCAACAGTTCAAGAGGAGCTCAGAACTTGCCCTCAACAGTGCAGCGCTTAGG 421
Db 121 GlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACACAGCTTGCAGCTGCGCTGCGAGTGCAGAGTGGCGGCGAGACAGATAC 472
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 157
```

RESULT 6

```
US-08-618-911-4
; Sequence 4, Application US/08618911
; Patent No. 5850016
; GENERAL INFORMATION:
; APPLICANT: Jung, Rudolf
; APPLICANT: Hastings, Craig
; APPLICANT: Coughlan, Sean
; APPLICANT: Hu, David
; TITLE OF INVENTION: ALTERATION OF AMINO ACID COMPOSITIONS IN
; TITLE OF INVENTION: SEEDS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: 700 Capital Square, 400 Locust Street
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/618,911
; FILING DATE: Concurrently herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Simon, Soma
; REGISTRATION NUMBER: 37,444
; REFERENCE/DOCKET NUMBER: 365-US
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-618-911-4

Alignment Scores:
Pred. No.: 9,89e-19 Length: 158
Score: 243.00 Matches: 56
Percent Similarity: 56.5% Conservative: 31
Best Local Similarity: 36.4% Mismatches: 55
Query Match: 19.0% Indels: 12
DB: 1 Gaps: 6

US-10-728-323-2 (1-717) x US-08-618-911-4 (1-158)

QY 5 ACATACATAGTAGCCTGCGCCCTTTCTCTCTGCTGCCACGATCTCGAGGAGCAG 64
Db 5 ThrIleLeuLeuIleSerLeuLeuPheCysIle---AlaHisThrCysSerAlaSerLys 23
QY 65 TGGGAATCCCAAGGAGACAGAGATGCCAGAGCCAGCTCGAGAGGCGCAACCTGAGGCC 124
Db 24 TrpGlnHisGlnAsp---SerCysArgLysGlnLeuGlnGlyValAsnLeuThrPro 42
QY 125 TCGAGAGCAACATCTCATGCAAGATCCAA-----CGTGACGAGGATTCATATGAACGG 178
Db 43 CysGluLysHisIleMetGluLysIleGlnGlyArgGlyAspAspAspAspAsp 62
QY 179 GACCGGTACAGCCCTAGTCAGGATCGGTACAGCCCTAGTCATATGATCGAGAGCGGCT 238
Db 63 AspAspAsnHisIleLeuArgThrMetArgGlyArgIleAsnTyrIleArgArgAsnGlu 82
QY 239 GGATCTCTCTCAG-----CACCAGAGAGGCTGTTCATAGCTGAGCTGAAC 280
Db 83 GlyLysAspGluAspGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 102
QY 281 GAGTTTGAGAACAAACAAAGGTGTCATGTCGAGGATTCGACACATATGAGAGAACCCAG 340
Db 103 GluLeu---ArgSerProLysCysGlnCysLysAlaLeuGlnLysIleMetGluAsnGln 121
QY 341 ACGGATAGTTTCAGGGGAGGCAACAGGAGCAACAGTTCACAGAGGAGGCTCAGGACTTG 400
Db 122 SerGluGluLeuGlu---GluLysGlnLysLysLysMetGluLysGluLeuIleAsnLeu 140
QY 401 CCTCAACAGTGGCGCTTAGGCGCACACAGCCCTTGGACTTG 442
Db 141 AlaThrMetCysArgPheGlyProMetIleGlnCysAspLeu 154

RESULT 7
US-08-938-675A-2
; Sequence 2, Application US/08938675A
; Patent No. 6107287
; GENERAL INFORMATION:
; APPLICANT: de Lumen, Benito O.
; APPLICANT: Galvez, Alfredo F.
; TITLE OF INVENTION: Lunasin Peptides
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
```


Db	42	ProCysGlu---HisIleMetGluLysaileGlnAlaGlyAArgArgGlyGluAspGlySer	50
Qy	173	GAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGA	232
Db	61	AspGluAspHisIleLeuIleAArgThrMetPro---GlyAArgIleAsnTrpIleArgLys	79
Qy	233	GGCGTCGATCCTCTCAG-----CACCAAGAGAGGTGTTGCAATCAGCTGAAC	280
Db	80	LysGluGlyLysGluGluGluGluGluGluGlyHisMetGlnLysCysCysSerGluMetSer	99
Qy	281	GAGTTTGAGAACCAACAAAGGTGCATGTGCGAGGCATTGCAACAGATCATGTCAGGAACCAAG	340
Db	100	GluLeuLys---SerProIleCysGlnCysIysAlaLeuGlnLysIleMetAspAsnGln	118
Qy	341	AGCGATAGTTGTCAGGGGACCAAGAGCAACAGTTCACAGAGGAGCTCAGGAACCTTG	400
Db	119	SerGlnGlnLeuGluGlyLys---GluLysLysGlnMetGluAArgGluLeuMetAsnLeu	137
Qy	401	CCTCAACAGTCGGCCCTTAGGGCACCAACAGCGTTTGCGACTTG	442
Db	138	AlaIleAArgCysAArgLeuGlyProMetIleGlyCysAspLeu	151

[illegible]

```

1 RESULT 10
2 US-08-618-911-6
3 ; Sequence 6, Application US/08618911
4 ; Patent No. 5850016
5 ;
6 ; GENERAL INFORMATION:
7 ; APPLICANT: Jung, Rudolf
8 ; APPLICANT: Hastings, Craig
9 ; APPLICANT: Coughlan, Sean
10 ; APPLICANT: Hu, David
11 ;
12 ; TITLE OF INVENTION: ALTERATION OF AMINO ACID COMPOSITIONS IN
13 ; SEEDS
14 ; NUMBER OF SEQUENCES: 13
15 ;
16 ; CORRESPONDENCE ADDRESS:
17 ; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
18 ; STREET: 700 Capital Square, 400 Locust Street
19 ; CITY: Des Moines
20 ; STATE: Iowa
21 ; COUNTRY: USA
22 ; ZIP: 50309
23 ;
24 ; COMPUTER READABLE FORM:
25 ; MEDIUM TYPE: Floppy disk
26 ; COMPUTER: IBM PC compatible
27 ; OPERATING SYSTEM: PC-DOS/MS-DOS
28 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
29 ; CURRENT APPLICATION DATA:
30 ; APPLICATION NUMBER: US/08/618,911
31 ; FILING DATE: Concurrently herewith
32 ; ATTORNEY/AGENT INFORMATION:
33 ; NAME: Simon, Soma
34 ;
35 ;
36 ;
37 ;
38 ;
39 ;
40 ;
41 ;
42 ;
43 ;
44 ;
45 ;
46 ;
47 ;
48 ;
49 ;
50 ;
51 ;
52 ;
53 ;
54 ;
55 ;
56 ;
57 ;
58 ;
59 ;
60 ;
61 ;
62 ;
63 ;
64 ;
65 ;
66 ;
67 ;
68 ;
69 ;
70 ;
71 ;
72 ;
73 ;
74 ;
75 ;
76 ;
77 ;
78 ;
79 ;
80 ;
81 ;
82 ;
83 ;
84 ;
85 ;
86 ;
87 ;
88 ;
89 ;
90 ;
91 ;
92 ;
93 ;
94 ;
95 ;
96 ;
97 ;
98 ;
99 ;
100 ;

```

REFERENCE/LOCAL NUMBER: 363-03
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 158 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-618-911-6

```

; MOLECULE TYPE: protein
US-08-618-911-6

Alignment Scores:
Pred. No.: 3,76e-17
Score: 229.00
Percent Similarity: 55.8%
Best Local Similarity: 35.1%
Query Match: 17.9%
DB: 1
Length: 158
Matches: 54
Conservative: 32
Mismatch: 56
Indels: 12
Gaps: 6

```

US-10-728-323-2 (1-717) x US-08-618-911-6 (1-158)


```
Db 21 ArgProCysGluGlnHisLeuMet 28
|||||
RESULT 12
US-08-670-186-4
; Sequence 4, Application US/08670186
; Patent No. 5859343
; GENERAL INFORMATION:
; APPLICANT: SUN, SAMUEL S.M.
; APPLICANT: XIONG, LIWEN
; APPLICANT: HU, ZHONG
; APPLICANT: CHEN, HANG
; TITLE OF INVENTION: RECOMBINANT SWEET PROTEIN MABINLIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 PENNSYLVANIA AVE NW, STE. 5500
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/670,186
; FILING DATE: 21-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 23461-20007.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 822-0168
; TELEX: 90-4030 MRSNFOERSWSH
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-670-186-4
Alignment Scores:
Pred. No.: 6,49e-09 Length: 158
Score: 156.00 Matches: 46
Percent Similarity: 47.6% Conservative: 32
Best Local Similarity: 28.0% Mismatches: 52
Query Match: 12.2% Indels: 34
DB: 1 Gaps: 8
US-10-728-323-2 (1-717) x US-08-670-186-4 (1-158)
QY 2 CTCACCATAGTAGCCCTTCCTCGCTGCTGCCACGCACTCGG---AGG 58
::: |||||
Db 5 IleLeuLeuLeuThrThrLeuAlaLeuPheValLeuAlaAsnAlaSerIleTyArg 24
::: |||||
QY 59 CAGCAGTGGGAACTC-----CAAGGAGACAGAAAGATGCCAGAGC 97
|||||
Db 25 ThrThrValGluLeuAspGluGluAspAsnAspGluAsnGlnProLeuCysArg 44
::: |||||
QY 98 CAGCTCGAGAGG---CCGAACCTGAGCCCTGGAGCAACATCTCATGCAAGATCCAA 154
||| ::::: |||||
Db 45 GlnPheGlnGlnHisGlnHisLeuArgAlaCysGlnArgTyIleArgArgAlaGln 64
||| ::::: |||||
QY 155 CGTGAC-----GAGGATTTCATATGAACGGGACCCGTCAGCCCTAGTCAGGATCCGTAC 208
||| |||
Db 65 ArgGlyGlyLeuValAspGluLeuGluLeuGluAspValGluGluAsnGluAsp----- 82
||| |||
QY 209 AGCCCTAGTCCATATGATCGGAGGCGCTGGATCCTCTCAGACCAACAGAGAGG----- 262
||| |||
```

```
Db 83 -----GluAspGluAsnGlnGlnArgGlyPro 91
|||||
QY 263 -----TGTTGCAATGAGCTGAACGAGTTTGACAAACAAAGGTCATGTGC 310
|||||
Db 92 AlaLeuArgLeuCysCysAsnGlnLeuArgGlnVal-----AsnLysProCysValCys 109
|||||
QY 311 GAGGCATTGCAACAGATCATGGAGAACCCAG-----AGCGATAGGTTGCGAGGGAGGCAA 364
|||||
Db 110 ProValLeuArgGlnAlaAlaHisGlnGlnLeuTyrglnGlnGlyGlnIleGluGlyProArg 129
|||||
QY 365 CAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTGCCTCAACAGTGCAGCCTTAGGGCA 424
|||||
Db 130 GlnValArgGlnLeuPheArgAlaAlaArgAsnLeuProAsnIleCysLysIleProAla 149
|||||
QY 425 CCACAGCGTTGC 436
|||||
Db 150 ValGlyArgCys 153
|||||
RESULT 13
US-08-670-186-6
; Sequence 6, Application US/08670186
; Patent No. 5859343
; GENERAL INFORMATION:
; APPLICANT: SUN, SAMUEL S.M.
; APPLICANT: XIONG, LIWEN
; APPLICANT: HU, ZHONG
; APPLICANT: CHEN, HANG
; TITLE OF INVENTION: RECOMBINANT SWEET PROTEIN MABINLIN
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 PENNSYLVANIA AVE NW, STE. 5500
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/670,186
; FILING DATE: 21-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 23461-20007.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 822-0168
; TELEX: 90-4030 MRSNFOERSWSH
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-670-186-6
Alignment Scores:
Pred. No.: 2,38e-08 Length: 158
Score: 151.00 Matches: 45
Percent Similarity: 47.0% Conservative: 32
Best Local Similarity: 27.4% Mismatches: 53
Query Match: 11.8% Indels: 34
DB: 1 Gaps: 8
US-10-728-323-2 (1-717) x US-08-670-186-6 (1-158)
QY 2 CTCACCATAGTAGCCCTTCCTCGCTGCTGCCACGCACTCGG---AGG 58
```

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-670-186-2

Alignment Scores:
Pred. No.: 5,34e-07 Length: 155
Score: 139.00 Matches: 45
Percent Similarity: 45.5% Conservative: 26
Best Local Similarity: 28.8% Mismatches: 47
Query Match: 10.9% Indels: 38
DB: 1 Gaps: 9

US-10-728-323-2 (1-717) x US-08-670-186-2 (1-155)

QY 5 ILELEULEUThrThrLeuAlaLeuPheValLeuAlaAsnAlaSerIleTyrArg 24
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 59 CAGCAGTGGGAAGTCTC-----CAAGGAGACAGAAGATGCCAGAGC 97
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 25 ThrThrValGluLeuAspGluGluAspAsnAspGluAsnGlnProLeuCysArgArg 44
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 98 CAGCTCGAGAGG---CGGAACCTGAGCCCTGCGAGCAACATCTCATGCAGAAGATCCAA 154
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 45 GlnPheGlnGlnHisGlnValArgAlaCysGlnArgTyrLeuArgArgArgAlaGln 64
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 155 CGTGAC-----GAGGATTCATATGAACGGGACCCCGTACAGCCCTAGTCAGGATCCGTAC 208
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 65 ArgGlyClyLeuAlaAspGluLeuGluLeuGluAspValGluGluAsnGluAsp----- 82
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 209 AGCCCTAGTCCATATGATCGAGAGCGGTGGATCTCTCAGCACCAAGAGAGG----- 262
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 83 -----GluAspGluAsnGlnArgGlyPro 91
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 263 -----TGTTGCAATGAGCTGAACGAGCTTTGAGAGCAACCAAGGTGCATGTGC 310
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 92 AlaLeuArgLeuCysCysAsnGlnLeuArgGlnVal-----AsnLysProCysValCys 109
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 311 GAGGCAATGCAACAGATCATGGAGAACCC-----AGCGATAGGTTGCAGGGGAGGCCAA 364
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 ProValLeuArgGlnAlaAlaHisGlnGlnLeuTyrGlnGlyGlnIleGluGlyProArg 129
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 365 CAGGAGCAACAGATTCAAGAGGGAGCTCAGGAACCTTGCCTCAACAGTGCAGCCCTTAGGGCA 424
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 130 GlnValArgArgLeuPheArgAlaAlaArgAsnLeuProAsnIleCysLysIleProAla 149
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 425 CCACAGGTTGC 436
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 150 ValGlyArgCys 153
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 14

```

US-08-670-186-2
; Sequence 2, Application US/08670186
; Patent No. 5859343
; GENERAL INFORMATION:
; APPLICANT: SUN, SAMUEL S.M.
; APPLICANT: XIONG, LIWEN
; APPLICANT: HU, ZHONG
; APPLICANT: CHEN, HANG
; TITLE OF INVENTION: RECOMBINANT SWEET PROTEIN MABINLIN
; CORRESPONDENCE ADDRESS:
; NUMBER OF SEQUENCES: 10
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 PENNSYLVANIA AVE NW, STE. 5500
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/670,186
; FILING DATE: 21-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 23461-20007.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 822-0168
; TELEX: 90-4030 MRSNFOERSWH
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 155 amino acids

```

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-670-186-2

Alignment Scores:
Pred. No.: 5,34e-07 Length: 155
Score: 139.00 Matches: 45
Percent Similarity: 45.5% Conservative: 26
Best Local Similarity: 28.8% Mismatches: 47
Query Match: 10.9% Indels: 38
DB: 1 Gaps: 9

US-10-728-323-2 (1-717) x US-08-670-186-2 (1-155)

QY 8 ATACTAGTAGCCCTCCCTTCTCTCTCGTCCCGCCAGCATCTCGGAGG----- 58
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 7 LeuPheAlaThrLeuAlaLeuPheValLeuLeuAlaAsnAlaSerIleGlnThrThrVal 26
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 59 -----CAGCAGTGGGAAGTCTC-----ArgCys 40
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 27 ValGluValAspGluGluAspAsnGlnLeuTrp-----ArgCys 40
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 92 CAGAGCAG---CTCGAGAGGGCGAACCTGAGGCCCTGCGAGCAACATCTCATGCAGAAG 148
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 41 GlnArgGlnPheLeuGlnHisGlnArgLeuArgAlaCysGlnArgPheIleHisArgArg 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 149 ATCCAA-----CGTGACGAGGATTCATATGAACGGGACCCGTACAGCCCTAGTCAGGAT 202
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 AlaGlnPheGlyGlyGlnProAspGluLeuGluAspGluValGluAspAsnAspAsp 80
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 203 CCGTACAGCCCTAGTCCATATGATCGAGAGGGCGTGGATCTCTCAGCACCAAGAGAGG 262
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 81 GluAsnGlnPro-----ArgArgProAla-----LeuArgGln 91
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 263 TGTTGCAATGAGCTGAACGAGTTTGAGAAACCAAGGTGCATGTGC----- 310
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 92 CysCysAsnGlnLeuArgGlnValAsp-----ArgProCysValCysProValLeuArg 109
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 311 GAGGCAATGCAACAGATCATGGAGAACCCAGAGCGATAGTTGAGGGGAGGCAACAGGAG 370
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 GlnAlaAlaGlnValLeuGlnArgGlnIle-----IleGlnGlyProGlnGlnLeu 127
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 371 CAACAGTTCAGAGGGGAGCTCAGGAACCTTGCCTCAACAGTGGCGGCTT 418
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 128 ArgArgLeuPheAspAlaAlaArgAsnLeuProAsnIleCysAsnIle 143
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 15
US-08-453-924-3
; Sequence 3, Application US/08453924
; Patent No. 580152
; GENERAL INFORMATION:
; APPLICANT: Kridl, Jean C.
; APPLICANT: Knauf, Vic C.
; TITLE OF INVENTION: Seed-Specific Transcriptional Regulation
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weil, Gotshal & Manges
; STREET: 2882 Sand Hill Road, Ste. 280
; CITY: Menlo Park
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,924
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

/ APPLICATION NUMBER: 07/742,834
/ FILING DATE: 08-AUGUST-1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/550,804
/ FILING DATE: 09-JULY-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/147,781
/ FILING DATE: 25-JANUARY-1988
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/078,538
/ FILING DATE: 28-JULY-1987
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 06/891,529
/ FILING DATE: 31-JULY-1986
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Rae-Venter, Barbara
/ REGISTRATION NUMBER: 32,750
/ REFERENCE/DOCKET NUMBER: CGNE-037/04US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 926-6200
/ TELEFAX: (415) 854-3713
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 184 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-453-924-3

Alignment Scores:
Pred. No.: 5.7e-07 Length: 184
Score: 139.00 Matches: 45
Percent Similarity: 41.8% Conservative: 29
Best Local Similarity: 25.4% Mismatches: 57
Query Match: 10.9% Indels: 46
DB: Gaps: 8

US-10-728-323-2 (1-717) x US-08-453-924-3 (1-184)

QY	8	ATACTAGTACGCTCGCCCTTTCTCTCGCTGCCCGCATCTGCG--AGGCAGCAG	64
Db	8	ValSerAlaThrLeuAlaLeuPheLeuThrAsnAlaSerValTyrArgThrVal	27
QY	65	TGGGAACCTCAAGGAGAC-----AGAAAGATGC	91
Db	28	ValGluValAspGluAspAlaThrAsnProAlaGlyProPheArgIleProLysCys	47
QY	92	CAGAGCCAGCTCAGAGGGCG--AACCTGAGCCCTGCGAGCAATCTCATGCAGAAG	148
Db	48	ArgLysGluPheGlnGlnAlaGlnHisLeuLysAlaCysGlnGlnTrpLeuHisLysGln	67
QY	149	ATCCAAACGTGAC-----GAGGATTCATATGAACGG	178
Db	68	AlaMetGlnSerGlySerGlyProSerTrpThrLeuAspGlyGluPheAspPheGluAsp	87
QY	179	GACCCGTACAGCCCTAGTCAGGATCGGTACAGCCCTAGTCCATATGATCGAGAGCGCT	238
Db	88	AspValGluAsnGlnGlnGlnGlyProGlnGlnArgProLeu-----	102
QY	239	GGATCTCTCAGACCACAGAGAGGTGTCATGAGCTGAACGAGTTTGAGAACACCAA	298
Db	103	-----LeuGlnGlnCysCysAsnGluLeuHisGlnGlu-----Pro	115
QY	299	AGTGTGATGTGCGAGCATTG-----CAACAGATCATGGAG	334
Db	116	LeuCysValCysProThrLeuLysGlyAlaSerLysAlaValLysGlnGlnIleArgGln	135
QY	335	AACCAAGCGATAGTTGAGGGAGGCAACAGGAGCAACAGTTTCAAGAGG-----	385
Db	136	GlnGlnGlyGlnMetGlnGlnGlnMetGlnGlnValIleSerArgIleTyrGln	155
QY	386	GAGCTCAGGAACCTTGCTCAACAGTCGGCGCTTAGGGCACACAGCGTTGC	436

Db 156 ThrAlaThrHisLeuProArgAlaCysAsnIleArgGlnValSerIleCys 172
Search completed: May 15, 2006, 22:28:14
JOB time : 23.562 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: May 15, 2006, 23:34:07 ; Search time 20.2364 Seconds
(without alignments)
4441.255 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 1281
Sequence: 1 gctcaccatactagtagcccc.....taaaagatcatgttttgg 717

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DSV=xlp
-Q=/abss/ABSSWEB_spool/US10728323/runat_15052006_172145_22607/app_query.fasta_1
-DB=Published Applications_AA_Main -QFMT=fastan -SUFFIX=n2p.rapbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blosum62 -TRANS=human0.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAESTZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02p
-USER=US10728323 @CGN_1_1_503 @runat_15052006_172145_22607 -NCPUP=6 -ICPU=3
-NO MMAP -NEG SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DBLOP=6 -DELEXT=7

Database : Published Applications_AA_Main.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	842	65.7	US-10-228-806-4
2	842	65.7	Sequence 4, Appli
3	838.5	65.5	US-10-100-303A-63
4	835	65.2	Sequence 2, Appli
5	819	63.9	US-10-899-551-4
6	819	63.9	Sequence 53, Appl
7	771	60.2	US-10-245-871-53
8	760	59.3	US-10-253-286-53
9	699	54.6	US-10-100-303A-82
10	463	36.1	US-10-899-551-56
11	243	19.0	US-10-731-221-77
			US-10-100-303A-81
			US-10-302-633-2

12	243	19.0	158	4	US-10-424-599-169225	Sequence 169225,
13	241.5	18.9	155	4	US-10-424-599-264372	Sequence 264372,
14	177.5	13.9	168	4	US-10-165-289A-2	Sequence 2, Appli
15	149.5	11.7	167	4	US-10-382-066-29	Sequence 29, Appl
16	144	11.2	169	4	US-10-165-289A-4	Sequence 4, Appli
17	140.5	11.0	165	4	US-10-425-115-190971	Sequence 190971,
18	139	10.9	184	3	US-09-782-130-3	Sequence 3, Appli
19	139	10.9	184	4	US-10-100-121-48	Sequence 48, Appl
20	129.5	10.1	150	4	US-10-442-174A-1	Sequence 1, Appli
21	127.5	10.0	151	4	US-10-424-599-167336	Sequence 167336,
22	127	9.9	22	4	US-10-228-806-67	Sequence 67, Appl
23	126	9.8	102	4	US-10-424-599-219338	Sequence 219338,
24	124	9.7	148	4	US-10-425-115-219350	Sequence 219350,
25	119	9.3	20	3	US-09-731-221-66	Sequence 66, Appl
26	118	9.2	261	4	US-10-291-265-715	Sequence 715, App
27	117	9.1	20	3	US-09-731-221-63	Sequence 63, Appl
28	116	9.1	20	3	US-09-731-221-60	Sequence 60, Appl
29	115	9.0	20	3	US-09-731-221-61	Sequence 61, Appl
30	113	8.8	20	3	US-09-731-221-65	Sequence 65, Appl
31	112.5	8.8	101	4	US-10-424-599-275916	Sequence 275916,
32	112	8.7	20	3	US-09-731-221-54	Sequence 54, Appl
33	112	8.7	20	3	US-09-731-221-59	Sequence 59, Appl
34	112	8.7	20	3	US-09-731-221-62	Sequence 62, Appl
35	112	8.7	20	3	US-09-731-221-64	Sequence 64, Appl
36	111	8.7	20	3	US-09-731-221-57	Sequence 57, Appl
37	111	8.7	20	3	US-09-731-221-74	Sequence 74, Appl
38	111	8.7	20	3	US-09-731-221-75	Sequence 75, Appl
39	111	8.7	20	3	US-09-731-221-76	Sequence 76, Appl
40	110	8.6	25	4	US-10-245-871-60	Sequence 60, Appl
41	110	8.6	25	4	US-10-253-286-60	Sequence 60, Appl
42	109	8.5	20	4	US-10-245-871-55	Sequence 55, Appl
43	109	8.5	20	4	US-10-253-286-55	Sequence 55, Appl
44	109	8.5	149	5	US-10-481-032A-244	Sequence 244, App
45	108	8.4	20	3	US-09-731-221-53	Sequence 53, Appl

ALIGNMENTS

RESULT 1
US-10-228-806-4
; Sequence 4, Application US/10228806
; Publication No. US20030049237A1
; GENERAL INFORMATION:
; APPLICANT: Bannont, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reactions
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0043
; CURRENT APPLICATION NUMBER: US/10/228,806
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-228-806-4

Alignment Scores:
Pred. No.: 1.22e-79 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 65.7% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-2 (1-717) x US-10-228-806-4 (1-157)
QY 2 CTCACCATACTAGTAGCCCTCGCCCTTTTCCTCGTCCGCCACGATCTGCGAGGCGAG 61
|||||
Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHisAlaSerAlaArgGln 20
|||||
QY 62 CAGTGGGAATCTCAAGGACAGAGATGCCAGCCACTCGAGAGGGCGAACCCTGAGG 121
|||||

Db 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCTGGAGCAACATCTCATGCAAGAGATCCAAAGTGCAGGAGATTATATGAACGGGAC 181
Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATATCGGAGAGCGCTGGA 241
Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 80
QY 242 TCCTCTCAGCACCAAGAGAGTGTTCGAATGAGCTGAACGAGTTTCAGAACCAACCAAGG 301
Db 81 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
QY 302 TGCATGTGCGAGGCATTGCAACAGATCATGCAAGATCCAAAGTGCAGGAGATTATATGAACGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAGTCCGCGCTTAGG 421
Db 121 GlnGlnGluGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACCACAGCTTGGACCTTGGACGTCGAAAGTGGCGGAGAGACAGATAC 472
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157
RESULT 2
US-10-100-303A-63
; Sequence 63, Application US/10100303A
; Publication No. US20030202980A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; FILE REFERENCE: to Allergy
; CURRENT APPLICATION NUMBER: US/10/100,303A
; CURRENT FILING DATE: 2002-03-18
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Arachis hypogaea, Prot/Nucleo Ara h 2
US-10-100-303A-63
Alignment Scores:
Pred. No.: 1.22e-79 Length: 157
Score: 842.00 Matches: 157
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 65.7% Indels: 0
DB: 4 Gaps: 0
US-10-728-323-2 (1-717) x US-10-100-303A-63 (1-157)
QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCATCTGCGAGGAGC 61
Db 1 LeuThrIleLeuValAlaLeuAlaLeuPheLeuAlaHisAlaSerAlaArgGln 20
QY 62 CAGTGGGAATCCAAAGAGACAGAAGATCCAGAGCCAGCTCGAGAGGGCGAACCTGAGG 121
Db 21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
QY 122 CCTGGAGCAACATCTCATGCAAGATCCAAAGTGCAGGAGATTATATGAACGGGAGC 181
Db 41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 60
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 241
Db 61 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 80
QY 242 TCCTCTCAGCACCAAGAGAGTGTTCGAATGAGCTGAACGAGTTTGAGAACCAACCAAGG 301

Db 81 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 100
QY 302 TGCATGTGCGAGGCATTGCAACAGATCATGCAAGATCCAAAGTGCAGGAGATTATATGAACGGGAGG 361
Db 101 CysMetCysGluAlaLeuGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 120
QY 362 CAACAGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAGTCCGCGCTTAGG 421
Db 121 GlnGlnGluGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 140
QY 422 GCACCACAGCTTGGACCTTGGACGTCGAAAGTGGCGGAGAGACAGATAC 472
Db 141 AlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArgTyr 157
RESULT 3
US-10-958-324-2
; Sequence 2, Application US/10958324
; Publication No. US20050114924A1
; GENERAL INFORMATION:
; APPLICANT: DODO, HORTENSE W.
; APPLICANT: ARNTZEN, CHARLES J.
; APPLICANT: KONAN, KOFFI N'DA
; APPLICANT: VIOUEZ, OLGA
; TITLE OF INVENTION: DOWN-REGULATION AND SILENCING OF ALLERGEN GENES IN
; FILE REFERENCE: 072121/0104
; CURRENT APPLICATION NUMBER: US/10/958,324
; CURRENT FILING DATE: 2004-10-06
; PRIOR APPLICATION NUMBER: US/09/715,036
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: 60/167,255
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-958-324-2
Alignment Scores:
Pred. No.: 3.08e-79 Length: 207
Score: 838.50 Matches: 162
Percent Similarity: 90.3% Conservative: 5
Best Local Similarity: 87.6% Mismatches: 17
Query Match: 65.5% Indels: 1
DB: 5 Gaps: 1
US-10-728-323-2 (1-717) x US-10-958-324-2 (1-207)
QY 2 CTCACCATAGTAGCCCTCGCCCTTTCTCTCGCTGCCACGCATCTGCGAGGAGC 61
Db 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuAlaHisAlaSerAlaArgGln 23
QY 62 CAGTGGGAATCCAAAGAGACAGAAGATCCAGAGCCAGCTCGAGAGGGCGAACCTGAGG 121
Db 24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
QY 122 CCTGGAGCAACATCTCATGCAAGATCCAAAGTGCAGGAGATTATATGAACGGGAGC 181
Db 44 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 63
QY 182 CCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGCGCTGGA 241
Db 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 83
QY 242 TCCTCTCAGCACCAAGAGAGTGTTCGAATGAGCTGAACGAGTTTGAGAACCAACCAAGG 301
Db 84 SerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 103
QY 302 TGCATGTGCGAGGCATTGCAACAGATCATGCAAGATCCAAAGTGCAGGAGATTATATGAACGGGAGG 361
Db 104 CysMetCysGluAlaLeuGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123

RESULT 5

```

US-10-245-871-53
: Sequence 53, Application US/10245871
: Publication No. US20030235594A1
: GENERAL INFORMATION:
: APPLICANT: HUMPHREYS, ROBERT
: APPLICANT: XU, MINZHEN
: TITLE OF INVENTION: II-KEY/ANTIGENIC EPTOPE HYBRID PEPTIDE VACCINES

```

```

? FILE REFERENCE: R8H-2013
? CURRENT APPLICATION NUMBER: US/10/245,871
? CURRENT FILING DATE: 2003-01-09
? PRIOR APPLICATION NUMBER: 10/197,000
? PRIOR FILING DATE: 2002-07-17
? PRIOR APPLICATION NUMBER: 09/396,813
? PRIOR FILING DATE: 1999-09-14
? NUMBER OF SEQ ID NOS: 905
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 53
? LENGTH: 156
? TYPE: PRT
? ORGANISM: Arachis hypogaea
? US-10-245-871-53

```

Alignment Scores:	3.26E-77	Length:	156
Pred. No.:	819.00	Matches:	153
Score:	100.0%	Conservative:	0
Percent Similarity:	100.0%	Mismatches:	0
Best Local Similarity:	100.0%	Indels:	0
Query Match:	63.9%	Gaps:	0
DB:	4		

US-10-728-323-2 (1-717) x US-10-245-871-53 (1-156)

Qy	2	CTCACCACTAGTAGCCCTCGCCCTTTTCTCTCTCGTGCACGCATCTCGAGGCAG	61
Db	4	LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaargGln	23
Qy	62	CAGTGGAACTCCAAGGAGCAGAGAAGATGCCAGAGCCAGCTCGAGAGGCGCAACCTTGAG	121
Db	24	GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuAArg	43
Qy	122	CCCTGCGAGCAACATCTCATGTCAGAAGATCCAACGTGACGAGGATTTCATATGAACGGGAC	181
Db	44	ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp	63
Qy	182	CCGTACAGCCCTAGTCAAGGATCCGTACAGCCCTAGTCCATATGATCGGAGAGGCCCTGCA	241
Db	64	ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly	83
Qy	242	TCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAACCAACCAAGG	301
Db	84	SerSerGlnHisGlnGluAArgCysCysAsnGluLeuAsnGluPheGluAsnGlnAArg	103
Qy	302	TGCATGTGCGAGGCAATTGCAACAGATCATGTAGAACCAGAGCGGATAGTTGACAGGGGAGG	361
Db	104	CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyAArg	123
Qy	362	CAACAGGAGCAACAGTTTCAAGAGGGAGCTCAGGAACCTTGCCCTCAACAGTGCGGCCCTTAGG	421
Db	124	GlnGlnGluGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuAArg	143
Qy	422	GCACCAAGCGTTCCGACTTGGAGCTGCAAAAGTGGCGGC	460
Db	144	AlaProGlnArgCysAspLeuAspValGluSerGlyGly	156

US-10-899-551-56

Alignment Scores:
Pred. No.: 5 58e-71 Length: 167
Score: 760.00 Matches: 139
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 59.3% Indels: 0
DB: 5 Gaps: 0

US-10-728-323-2 (1-717) x US-10-899-551-56 (1-167)

QY 53 GCGAGCAGAGTGGGAATCTCAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGGCG 112
Db 17 AlaaArgGlnGlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAla 36
QY 113 AACCTGAGCCCTCGCAGCAACATCTCATGCAGAAATCCACGTCACGAGGATTCATAT 172
Db 37 AsnLeuArgProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyr 56
QY 173 GAACGGACCCGACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATATCGGAGA 232
Db 57 GluArgAspProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArg 76
QY 233 GCGCTGGATCTCTTCAGCACCAAGAGAGGTGTTCGAATCAGCTGAACGAGTTTGAGAAC 292
Db 77 GlyAlaGlySerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGluAsn 96
QY 293 AACCAAGTGCATGTCGAGGATTCGACAGATTCGACAGATCATGCAGACGAGCCGATAGTTG 352
Db 97 AsnGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeu 116
QY 353 CAGGGAGGCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTGCTCAACAGTGC 412
Db 117 GlnGlyArgGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCys 136
QY 413 GCGCTTAGGCAACACAGCGTTGCGACTTGGAGCTCGAAAGTGGCGGACAGACAGA 469
Db 137 GlyLeuArgAlaProGlnArgCysAspLeuAspValGluSerGlyGlyArgAspArg 155

RESULT 9

US-09-731-221-77
; Sequence 77, Application US/09731221
; Patent No. US20020018778A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael
; TITLE OF INVENTION: Passive Desensitization
; FILE REFERENCE: 2002834-0103
; CURRENT APPLICATION NUMBER: US/09/731,221
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 77
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Arachis
; OTHER INFORMATION: Hypogaea
US-09-731-221-77

Alignment Scores:
Pred. No.: 1 52e-64 Length: 166
Score: 699.00 Matches: 131
Percent Similarity: 92.9% Conservative: 0
Best Local Similarity: 92.9% Mismatches: 10
Query Match: 54.6% Indels: 0
DB: 3 Gaps: 0

US-10-728-323-2 (1-717) x US-09-731-221-77 (1-166)

QY 50 TTCGCGAGCAGAGTGGGAATCTCAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGG 109
|||||

Db 15 SerAlaArgGlnGlnAlaGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuAlaArg 34
QY 110 GCGAACCTGAGGCCCTCGCAGCAACATCTCATGCAGAAAGATCCAACTGACGAGGATTC 169
Db 35 AlaAsnLeuArgAlaCysGluAlaHisLeuMetGlnLysIleGlnAlaAspGluAspSer 54
QY 170 TATGAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATATCGG 229
Db 55 TyrGluArgAlaProTyrSerProSerGlnAlaProTyrSerProSerProTyrAspArg 74
QY 230 AGAGCGGCTGGATCTCTTCAGCACCAAGAGAGGTGTTCGAATCAGCTGACGAGTTGAG 289
Db 75 ArgGlyAlaGlySerSerGlnHisGlnGluArgCysAsnGluLeuAsnGluPheGlu 94
QY 290 AACCAACCAAGGTGATGTCGAGGATTCGAAACAGATCATGCAGAACCCAGAGCCGATAGG 349
Db 95 AsnAsnGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArg 114
QY 350 TTGAGGCGGAGGCAACAGGAGCAACAGTTCAAGAGGAGCTCAGGAACCTGCTCAACAG 409
Db 115 LeuGlnGlyAlaGlnGlnGlnPheLysArgGluAlaArgAsnLeuProGlnGln 134
QY 410 TCGGCGCTTAGGCGACACAGCGTTGCGACTTGGAGCTCGAAAGTGGCGGACAGACAGA 469
Db 135 CysGlyLeuArgAlaProGlnArgCysAspAlaAspValGluSerGlyGlyArgAspArg 154
QY 470 TAC 472
|||
Db 155 Tyr 155

RESULT 10

US-10-100-303A-81
; Sequence 81, Application US/10100303A
; Publication No. US20030202980A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; FILE REFERENCE: to Allergy
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100,303A
; CURRENT FILING DATE: 2002-03-18
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 81
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Arachis hypogaea, Prot/Nucleo Ara h 2
US-10-100-303A-81

Alignment Scores:
Pred. No.: 1e-39 Length: 83
Score: 463.00 Matches: 83
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 36.1% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-2 (1-717) x US-10-100-303A-81 (1-83)

QY 68 GAACTCCAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGCGCAACCTGAGGCCCTGC 127
Db 1 GluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArgProCys 20
QY 128 GAGCAACATCTCATGCAGAAATCCAACTGACGAGGATTCATATGAACGGGACCGTAC 187
Db 21 GluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAspProTyr 40
QY 188 AGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGGCGCTGCATCCT 247
Db 41 SerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGlySerSer 60
QY 248 CAGCACCAAGAGAGGTGTTCGAATCAGCTGAACGAGTTTGAGAAACAACCAAGGTCATG 307
|||||

Db 61 GlnHisGlnArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArgCysMet 80
QY 308 TGGAGGCA 316
|||||
Db 81 CysGluAla 83

RESULT 11
US-10-302-633-2
; Sequence 2, Application US/10302633
; Publication No. US20030229038A1
; GENERAL INFORMATION:
; APPLICANT: de Lumen, Benito O.
; TITLE OF INVENTION: Lunasin Peptides
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/302,633
; FILING DATE: 22-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/531,727
; FILING DATE: 21-Mar-2000
; APPLICATION NUMBER: 08/938,675
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B98-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-302-633-2

Alignment Scores:
Pred. No.: 1.94e-16 Length: 158
Score: 243.00 Matches: 56
Percent Similarity: 56.5% Conservative: 31
Best Local Similarity: 36.4% Mismatches: 55
Query Match: 19.0% Indels: 12
DB: 4 Gaps: 6

US-10-728-323-2 (1-717) x US-10-302-633-2 (1-158)
QY 5 ACCATACTAGTAGCCCTCGCCCTTTCTCTCTCGTCCACGATCTGGAGGACGAG 64
|||||
Db 5 ThrileLeuLeuIleSerLeuPheCysile---AlaHisThrCysSerAlaSerLys 23
|||||
QY 65 TGGGAACCTCCAGAGACAGAGATGCCAGAGCCAGCTCGAGGGCGGACCTGAGGCC 124
|||||
Db 24 TrpGlnHisGlnAsp---SerCysArgLysGlnLeuGlnGlyValAsnLeuThrPro 42
|||||
QY 125 TGGCAGCAACATCTCATGCAGAGATCCAA-----CGTGACGAGGATTCATATGACGG 178
|||||
Db 5 ThrileLeuLeuIleSerLeuPheCysile---AlaHisThrCysSerAlaSerLys 23
|||||
QY 65 TGGGAACCTCCAGAGACAGAGATGCCAGAGCCAGCTCGAGGGCGGACCTGAGGCC 124
|||||
Db 24 TrpGlnHisGlnAsp---SerCysArgLysGlnLeuGlnGlyValAsnLeuThrPro 42
|||||
QY 125 TGGCAGCAACATCTCATGCAGAGATCCAA-----CGTGACGAGGATTCATATGACGG 178
|||||

Db 43 CysGluLysHisIleMetGluLysIleGlnGlyArgGlyAspAspAspAspAsp 62
QY 179 GACCCGTACAGCCCTAGTCAGATCGGTACAGCCCTAGTCCATATGATCGGAGCGCT 238
|||||
Db 63 AspAspAsnHisIleLeuArgThrMetArgGlyArgIleAsnTyrlleArgArgAsnGlu 82
|||||
QY 239 GGATCCTCTCAG-----CACCAGAGAGGTGTTCACATAGCTGAAC 280
|||||
Db 83 GlyLysAspGluAspGluGluGluGlyHisMetGlnLysCysCysThrGluMetSer 102
|||||
QY 281 GAGTTTGTGAGAACCAACAAAGTTCATGTCGAGGATTCACACAGATCATGGAGAACCCAG 340
|||||
Db 103 GluLeu---ArgSerProLysCysGlnCysLysAlaLeuGlnLysIleMetGluAsnGln 121
|||||
QY 341 AGCGATAGTTTCAGGGGAGGCAACAGAGGACCAACAGTTCAGAGGAGGCTCAGGAACCTG 400
|||||
Db 122 SerGluGluLeuGlu---GluLysGlnLysLysLysMetGluLysGluLeuIleAsnLeu 140
|||||
QY 401 CCTCAACAGTGGCGCTTAGGCGCACACAGCGTTGCGACTTG 442
|||||
Db 141 AlaThrMetCysArgPheGlyProMetIleGlnCysAspLeu 154

RESULT 12
US-10-424-599-169225
; Sequence 169225, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 169225
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_123825C.1.pap
US-10-424-599-169225

Alignment Scores:
Pred. No.: 1.94e-16 Length: 158
Score: 243.00 Matches: 56
Percent Similarity: 56.5% Conservative: 31
Best Local Similarity: 36.4% Mismatches: 55
Query Match: 19.0% Indels: 12
DB: 4 Gaps: 6

US-10-728-323-2 (1-717) x US-10-424-599-169225 (1-158)
QY 5 ACCATACTAGTAGCCCTCGCCCTTTCTCTCTCGTCCACGATCTGGAGGACGAG 64
|||||
Db 5 ThrileLeuLeuIleSerLeuPheCysile---AlaHisThrCysSerAlaSerLys 23
|||||
QY 65 TGGGAACCTCCAGAGACAGAGATGCCAGAGCCAGCTCGAGGGCGGACCTGAGGCC 124
|||||
Db 24 TrpGlnHisGlnAsp---SerCysArgLysGlnLeuGlnGlyValAsnLeuThrPro 42
|||||
QY 125 TGGCAGCAACATCTCATGCAGAGATCCAA-----CGTGACGAGGATTCATATGACGG 178
|||||
Db 43 CysGluLysHisIleMetGluLysIleGlnGlyArgGlyAspAspAspAspAsp 62
|||||
QY 179 GACCCGTACAGCCCTAGTCAGATCGGTACAGCCCTAGTCCATATGATCGGAGCGCT 238
|||||
Db 63 AspAspAsnHisIleLeuArgThrMetArgGlyArgIleAsnTyrlleArgArgAsnGlu 82
|||||
QY 239 GGATCCTCTCAG-----CACCAGAGAGGTGTTCACATAGCTGAAC 280
|||||

Search completed: May 16, 2006, 00:00:30
Job time : 104.182 secs

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 23:41:32 ; Search time 2.46663 Seconds
(without alignments)
4094.130 Million cell updates/sec

Title: US-10-728-323-2

Perfect score: 1281

Sequence: 1 gctcaccatactagtagccccc.....taaaagatcatgttttgg 717

Scoring table: BLASTUM62
Xgapop 10.0 , Xgapext 0.5
Ygapcp 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 250354 seqs, 46948837 residues

Total number of hits satisfying chosen parameters: 500708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/abs/ABSWEB/spool/US10728323/runat_15052006_172148_22650/app_query.fasta.1
-DB=Published Applications AA New -GPM=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abs02p
-USER=US10728323 @CN 1.1.63 @runat_15052006_172148_22650 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DSEXT=7

Database : Published Applications AA New.*

1: /SID55/ptodata/1/pubpaa/US08 NEW PUB.pep1.*
2: /SID55/ptodata/1/pubpaa/US06 NEW PUB.pep1.*
3: /SID55/ptodata/1/pubpaa/US07 NEW PUB.pep1.*
4: /SID55/ptodata/1/pubpaa/US08 NEW PUB.pep1.*
5: /SID55/ptodata/1/pubpaa/PCT NEW PUB.pep1.*
6: /SID55/ptodata/1/pubpaa/US09 NEW PUB.pep1.*
7: /SID55/ptodata/1/pubpaa/US09 NEW PUB.pep1.*
8: /SID55/ptodata/1/pubpaa/US10 NEW PUB.pep1.*
9: /SID55/ptodata/1/pubpaa/US10 NEW PUB.pep1.*
10: /SID55/ptodata/1/pubpaa/US11 NEW PUB.pep1.*
11: /SID55/ptodata/1/pubpaa/US11 NEW PUB.pep1.*
12: /SID55/ptodata/1/pubpaa/US60 NEW PUB.pep1.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	819	63.9	156	11	US-11-033-039-53
2	243	19.0	9	9	Sequence 53, Appl
3	241.5	18.9	155	9	US-10-508-263-44
4	184.5	14.4	323	9	Sequence 42, Appl
5	151	11.8	295	9	US-10-508-263-52
					Sequence 52, Appl
					Sequence 50, Appl

6	146.5	11.4	164	9	US-10-508-263-2
7	118	9.2	261	11	US-11-000-463-715
8	116.5	9.1	319	11	US-11-096-568A-24535
9	110	8.6	25	11	US-11-033-039-60
10	109	8.5	20	11	US-11-033-039-55
11	107.5	8.4	260	11	US-11-030-908-7
12	106.5	8.3	178	9	US-10-508-263-32
13	105.5	8.2	164	9	US-10-508-263-4
14	105.5	8.2	180	9	US-10-508-263-38
15	104	8.1	144	9	US-10-508-263-40
16	103.5	8.1	166	9	US-10-508-263-8
17	103.5	8.1	178	9	US-10-508-263-34
18	102.5	8.0	178	9	US-10-508-263-46
19	101.5	7.9	178	9	US-10-508-263-36
20	99.5	7.8	148	9	US-10-526-716-2
21	96	7.5	144	9	US-10-508-263-48
22	95	7.4	170	9	US-10-508-263-6
23	95	7.4	568	9	US-10-784-004-390
24	92	7.2	336	11	US-11-172-740-126
25	91	7.1	297	9	US-10-203-486-3
26	90	7.0	450	11	US-11-119-569-21
27	90	7.0	448	11	US-11-119-569-4
28	89	6.9	1341	11	US-11-188-298-21361
29	87.5	6.8	368	11	US-11-188-298-13415
30	87	6.8	585	9	US-10-784-004-712
31	87	6.8	1236	9	US-10-873-528-109
32	86.5	6.8	628	11	US-11-183-136-10
33	86	6.7	1230	11	US-11-087-099-1702
34	86	6.7	1230	11	US-11-087-099-8922
35	86	6.7	1230	11	US-11-188-298-1685
36	86	6.7	1230	11	US-11-188-298-8275
37	85.5	6.7	185	11	US-11-264-096-1711
38	85.5	6.7	378	11	US-11-172-740-1335
39	85.5	6.7	638	11	US-11-072-512-2626
40	85.5	6.7	1162	10	US-11-106-014-94
41	85	6.6	15	11	US-11-033-039-54
42	85	6.6	20	11	US-11-033-039-59
43	84.5	6.6	280	11	US-11-072-512-3063
44	84.5	6.6	716	11	US-11-194-890-20
45	83.5	6.5	289	11	US-11-096-568A-8938

ALIGNMENTS

RESULT 1
US-11-033-039-53
; Sequence 53, Application US/11033039
; Publication No. US20060002947A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: LI-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2017US01
; CURRENT APPLICATION NUMBER: US/11/033,039,
; CURRENT FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: 10/245,871
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 1452
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 53
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-11-033-039-53

Alignment Scores:
Pred. No.: 2.01e-73 Length: 156
Score: 819.00 Matches: 153
Percent Similarity: 100.0% Conservative: 0

```

Best Local Similarity: 100.0%      Mismatches: 0
Query Match: 63.9%               Indels: 0
DB: 11                          Gaps: 0

US-10-728-323-2 (1-717) x US-11-033-039-53 (1-156)

QY 2 CTCACCATAGTAGCCCTCGCCCTTTCCTCTCGCTGCCACGCATCTGCGAGCAG 61
Db 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHisSerAlaArgGln 23
QY 62 CAGTGGAACTCAAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAGG 121
Db 24 GlnTrpGluLeuGlnGlyAspArgAcGysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
QY 122 CCCTGCGGACCAATCTTCATGCAAGATCCACGTCGACGAGATTTCATATGAAACGGGAC 181
Db 44 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 63
QY 182 CCCTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGGCGCTGGA 241
Db 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrSerProTyrSerProTyr 83
QY 242 TCCTCTCAGACCAAGAGAGGTGTGCAATGAGCTGAACGAGTTTGAGAACCAACCAAGG 301
Db 84 SerSerGlnHisGlnGluArgCysCysAsnGluLeuAsnGluPheGluAsnAsnGlnArg 103
QY 302 TGCATGTGCGAGCATTTGCAACAGATCATGGAGAACAGAGCCGATAGTTGCGAGGGAGG 361
Db 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 123
QY 362 CACAGGAGCAACAGTTCAGAGGAGGCTCAGAACTTGCTCAACAGTCCGCGCTTAGG 421
Db 124 GlnGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
QY 422 GCACCAAGGCTTGGCAGCTTGGAGCTGCAAGTGGCGGC 460
Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyGly 156

RESULT 2
US-10-508-263-44
; Sequence 44, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Glycine max
US-10-508-263-44

Alignment Scores:
Pred. No.: 4,06e-16      Length: 158
Score: 243.00           Matches: 56
Percent Similarity: 56.5%      Conservative: 31
Best Local Similarity: 36.4%      Mismatches: 55
Query Match: 19.0%          Indels: 12
DB: 9                  Gaps: 6

US-10-728-323-2 (1-717) x US-10-508-263-44 (1-158)

QY 5 ACCATAGTAGTAGCCCTCGCCCTTTCCTCTCGCTGCCACGCATCTGCGAGCAGCAG 64
Db 5 ThrIleLeuLeuSerLeuLeuPheCysIle---AlaHisThrCysSerAlaSerLys 23
QY 65 TGGGAACCTCAAGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGGCGAACCTGAGGCC 124
Db 24 TrpGlnHisGlnGlnAsp---SerCysArgLysGlnLeuGlnGlyValAsnLeuThrPro 42

```

```

QY 125 TCGGAGCAACATCTCATGTCAGAGATCCAA-----CGTGACGAGATTTCATATGAACGG 178
Db 43 CysGluLysHisIleMetGluLysIleGlnGlyArgGlyAspAspAspAspAspAsp 62
QY 179 GACCCGTACAGCCCTAGTCAGATCGGTACAGCCCTAGTCCATATGATCGGAGAGCGCT 238
Db 63 AspAspAsnHisIleLeuArgThrMetArgGlyArgIleAsnTyrIleArgArgAsnGlu 82
QY 239 GGATCCTCTTCAG-----CACCAAGAGAGGTGTTCATGAGTGAAC 280
Db 83 GlyLysAspGluAspGluGluGluGlyHisMetGlnLysCysCysThrGluMetSer 102
QY 281 GAGTTTGAGAACCAACAAAGGTGCATGTGGAGGCATTCACACAGATCATGGAGAACCCAG 340
Db 103 GluLeu---ArgSerProLysCysGlnCysLysAlaLeuGlnLysIleMetGluAsnGln 121
QY 341 AGCGATAGGTTCAGGGGAGGCAACAGAGAGCAACAGTTTCAAGAGAGGAGCTCAGAACTTG 400
Db 122 SerGluGluLeuGlu---GluLysGlnLysLysLysMetGluLysGluLeuIleAsnLeu 140
QY 401 CCTCAACAGTTCGGCCCTTAGGGCACACAGCGTTTGGACTTG 442
Db 141 AlaThrMetCysArgPheGlyProMetIleGlnCysAspLeu 154

RESULT 3
US-10-508-263-42
; Sequence 42, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Glycine max
US-10-508-263-42

Alignment Scores:
Pred. No.: 5.7e-16      Length: 155
Score: 241.50           Matches: 58
Percent Similarity: 57.1%      Conservative: 30
Best Local Similarity: 37.7%      Mismatches: 53
Query Match: 18.9%          Indels: 13
DB: 9                  Gaps: 7

US-10-728-323-2 (1-717) x US-10-508-263-42 (1-155)

QY 2 CTCACCATAGTAGCCCTCGCCCTTTCCTCTCGCTGCCACGCATCTGCGAGCAG 61
Db 4 LeuThrIleLeuLeuIleAlaLeuLeuPheIle-----AlaHisThrCysCysAlaSer 21
QY 62 CAGTGGAACTCAAGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGGGCGAACCTGAGG 121
Db 22 LysTrpGlnGlnHisGlnGlnGluSerCysArgGluGlnLeuLysGlyIleAsnLeuAsn 41
QY 122 CCCTGCGGACCAACATCTTCATGCAAGATCCAA-----CGTGACGAGATTTCATAT 172
Db 42 ProCysGlu---HisIleMetGluLysIleGlnAlaGlyArgArgGlyGluAspGlySer 60
QY 173 GAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGATCGGAGA 232
Db 61 AspGluAspHisIleLeuIleArgThrMetPro---GlyArgIleAsnTyrIleArgLys 79
QY 233 GCGCGTGGATCCTCTCAG-----CACCAAGAGAGGTGTTCATGAGTGAAC 280
Db 80 LysGluGlyLysGluGluGluGluGluGlyHisMetGlnLysCysCysSerGluMetSer 99

```



```

; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Zhou, Ping
; APPLICANT: Cao, Yi-Cheng
; APPLICANT: Drmanac, Radjic T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 785CIP4CN
; CURRENT APPLICATION NUMBER: US/11/000,463
; CURRENT FILING DATE: 2004-11-29
; PRIOR APPLICATION NUMBER: 10/291,265
; PRIOR FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: PCT/US01/02623
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/922,279
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 715
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-000-463-715

Alignment Scores:
Pred. No.: 0.00122 Length: 261
Score: 118.00 Matches: 47
Percent Similarity: 51.1% Conservative: 23
Best Local Similarity: 34.3% Mismatches: 61
Query Match: 9.2% Indels: 6
DB: 11 Gaps: 4

US-10-728-323-2 (1-717) x US-11-000-463-715 (1-261)
QY 460 GCGCGCACTTTCGAGCTCAAGTCGCAACGCTGTGTGGTCCCTAAGCGCGCACATGTTGAGG 401
||||| : : : : : ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12 AlaAlaLeuValLysMetGluLeuValArgTrpAlaProValAlaCysLeuValArg 31
QY 400 CAAGTTCCTGAGCTCCCTCTTGAACATGTTGTCTCTGTTGCTCCCTCCCTGCAACCATATCGCT 341
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 32 GluValAlaLeuGluProLeuAlaLeuValLeuValGluMetMet-----ValLeu 49
QY 340 CTGGTTCTCCATGATCTGTTGCAATGCTCGCATGCTCCCATGACCTTTGGTGTCTCTC---AAA 284
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 50 LeuValLeuProGlyProLeu-----ValProAlaProLeuValLeuLeuAlaSer 67
QY 283 CTCGTTGAGCTCATTGCAACACCTCTCTTGGTGTGTGAGAGGATCCAGCGCTCTCCGATC 224
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 68 LeuValLeuValLeuArgValLysLeuValProLysGlyProGluAlaLeuLysVal 87
QY 223 ATATGGACTAGGCGGTGTACGGATCTGACTAGGCGCTGTACGGGTCCCGTTCATATGATC 164
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 88 ProArgValCysValValSerLeuAlaProLeuAlaLeuVal---LeuLeuAlaLeu 106
QY 163 CTCGTCAGCTGGATCTTCTGCATGAGATGTGCTCGCAGGCGCTCAGGTTGCGCCCTCTC 104
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 107 LeuGluThrLeuValLeuArgGluSerLeuValLeuLysValProMetValLeuLeuVal 126
QY 103 GAGCTGGCTCTGGCATCTCTGTCTCTCTGAGTTCACCTGTGCTCGC 53
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127 LeuLeuValLeuAlaSerLeuValProGluAlaProLeuAspProArg 143

RESULT 8
US-11-096-568A-24535
; Sequence 24535, Application US/11096568A
; Publication No. US20060048240A1

```

```
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: Therby
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 24535
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(319)
; OTHER INFORMATION: Ceres Seq. ID no. 12435624
US-11-096-568A-24535

Alignment Scores:
Pred. No.: 0.0018 Length: 319
Score: 116.50 Matches: 53
Percent Similarity: 37.7% Conservative: 16
Best Local Similarity: 29.0% Mismatches: 69
Query Match: 9.1% Indels: 45
DB: 11 Gaps: 8

US-10-728-323-2 (1-717) x US-11-096-568A-24535 (1-319)
QY 1 GTCACCATAGTAGAGCTGCGCTTTTCCTCTGCTGCCAGCGATCTGCAGGCA 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 146 AlaProSerAlaSerAlaProThrProCysProSerThrSerProArgAlaCysArgSer 165
QY 61 CGAGTGGGAACCT--CCAGGAGACAGAAGATGCCAGAGCCAGCTCGAGAGGCGCAACT 117
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 166 AlaArgGlySerAlaProArgProProProProProProProProProSerThrPro 185
QY 118 GAGGCCCTCGCAACATCTCATGCGAAGATCCAACTGAGGAGGATTCATATGAAG 177
   ::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 186 SerSerAlaArgAla--SerAlaAlaThrThrSerSerSerArgGly-SerSerArgAr 204
QY 178 GGACCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCATATGATCGGAGAGGCC 237
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 204 gargProSerAlaAlaSerThrProThrThrSerProProProSerSerAlaAlaSerSe 224
QY 238 TGGATCCTCTCAGACC-----AAGAGAGGTG 264
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 224 rSerSerAla-AlaThrThrProPheThrSerSerArgPheProSerArgGlnArgSerA 244
QY 265 TTGCAATGAGCTGAACGAGT-----TTGAGAACAAACCAAGGTGTCATGTGCAGGC 315
   ::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 244 laSerThrSerSerTrpSerProThrSerArgProArgAlaArgCysAlaProA 264
QY 316 ATGC-----AACAGAT 327
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 264 rgCysProGlyArgSerThrSerSerThrCysAlaThrProAlaArgGlnArgArGs 284
QY 328 CATGGAGACCAAGCAGTAGGTGCGAGGGGCAACAGGACACAGTTCACAGGGA 387
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 284 erTrpArg-----CysCysArgGlyThrArgAlaSerSerAlaProArg---- 299
QY 388 GCTCAGGAATTCCTCAACAGTGGCGCTTA-----GGGACCAACAGCGTTGCCA 438
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 300 -----CysArgProThrAlaSerTrpSerProProGlyHisProSerTrpLeua 316
QY 439 CTTGG 443
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 316 laTrp 317

RESULT 9
US-11-033-039-60
; Sequence 60, Application US/11033039
; Publication No. US2006002947A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: LI-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2017US01
; CURRENT APPLICATION NUMBER: US/11/033,039
; CURRENT FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: 10/245,871
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 1452
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 55
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-11-033-039-55

Alignment Scores:
Pred. No.: 0.00576 Length: 20
Score: 109.00 Matches: 20
```

```
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: LI-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2017US01
; CURRENT APPLICATION NUMBER: US/11/033,039
; CURRENT FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: 10/245,871
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 1452
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 60
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: hybrid peptide
; FEATURE:
; NAME/KEY: MOD.RES
; LOCATION: (5)..(5)
; OTHER INFORMATION: Ava
US-11-033-039-60

Alignment Scores:
Pred. No.: 0.00479 Length: 25
Score: 110.00 Matches: 21
Percent Similarity: 84.0% Conservative: 0
Best Local Similarity: 84.0% Mismatches: 4
Query Match: 8.6% Indels: 0
DB: 11 Gaps: 0

US-10-728-323-2 (1-717) x US-11-033-039-60 (1-25)
QY 101 CTCGAGAGGGCGAACCCTGAGCCCTGCGAGCAACATCTCATGCAAGATCCAACGTGAC 160
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 LeuArgMetLys**LeuArgProCysGluGlnHisLeuMetGlnLysIleGlnArgasp 20
QY 161 GAGGATTCATATGAA 175
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21 GluAspSerTyrglu 25

RESULT 10
US-11-033-039-55
; Sequence 55, Application US/11033039
; Publication No. US2006002947A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: LI-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2017US01
; CURRENT APPLICATION NUMBER: US/11/033,039
; CURRENT FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: 10/245,871
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 1452
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 55
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-11-033-039-55

Alignment Scores:
Pred. No.: 0.00576 Length: 20
Score: 109.00 Matches: 20
```

Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 8.5% Indels: 0
 DB: 11 Gaps: 0

US-10-728-323-2 (1-1717) x US-11-033-039-55 (1-20)

QY 116 CTGAGCCCTCGAGCAACATCTCATGCGAGAGATCAAGCTGACGAGGATTCATATGAA 175
 Db 1 LeuArgProCysGluGlnHisLeuMetGlnLysLeuGlnArgAspGluAspSerTyrGlu 20

RESULT 11

US-11-090-908-7
 ; Sequence 7, Application US/11090908
 ; Publication No. US20050255574A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pedersen, Henrik
 ; APPLICANT: Poulsen, Charlotte Horsmans
 ; APPLICANT: Soe, Jorn Borch
 ; APPLICANT: Zargahi, Masoud Rajabi
 ; TITLE OF INVENTION: Proteins
 ; FILE REFERENCE: 674509-2011.1
 ; CURRENT APPLICATION NUMBER: US/11/090,908
 ; CURRENT FILING DATE: 2005-03-25
 ; PRIORITY APPLICATION NUMBER: PCT/IB2003/004585
 ; PRIOR FILING DATE: 2003-09-26
 ; PRIOR APPLICATION NUMBER: GB 0222512.6
 ; PRIOR FILING DATE: 2002-09-27
 ; PRIOR APPLICATION NUMBER: GB 0223674.3
 ; PRIOR FILING DATE: 2002-10-10
 ; PRIOR APPLICATION NUMBER: GB 0228082.4
 ; PRIOR FILING DATE: 2002-12-02
 ; PRIOR APPLICATION NUMBER: US 60/431,209
 ; PRIOR FILING DATE: 2002-12-05
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 7
 ; LENGTH: 260
 ; TYPE: PRT
 ; ORGANISM: Triticum sp.

US-11-090-908-7

Alignment Scores:
 Pred. No.: 0.0136 Length: 260
 Score: 107.50 Matches: 44
 Percent Similarity: 40.9% Conservative: 21
 Best Local Similarity: 27.7% Mismatches: 60
 Query Match: 8.4% Indels: 34
 DB: 11 Gaps: 9

US-10-728-323-2 (1-1717) x US-11-090-908-7 (1-260)

QY 1 GCTCACCATACTAGT-----AGCCCTCGCCCTTTCTCTCTCTCTGC----- 42
 Db 102 AlaHisLeuLysSerMetArgMetSerLeuGlnThrLeuProSerMetCysAsnIleTyr 121
 QY 43 ---CCAGCATCTCGAGGAGAGAGTGGGAATCCCAAGGAGACAGAGAAGATCCAGAGCCCA 99
 Db 122 ValProValGlnCysGln-GlnGlnGlnGlnLeu-----GlyArgGlnGlnGlnGln 139
 QY 100 GCTCGAGAGGGCAACCTGAGGCCCTCGAGCAACATCTCATGCGAAGATCCAAACGTGA 159
 Db 139 nLeuGlnGlu---GlnLeuLysProCysAlaThrPheLeuGlnHisGlnCysArgProMe 158
 QY 160 CGAGGATTCATATGAACGGGACCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCC 219
 Db 158 tThrValProPhe-----ProHisThrProValGlnLysProThrSer----- 172
 QY 220 ATATGATCGGAGAGGCCCTGATCCTCTCAGCACCAGAGAGGTGTTCGAATGAGCTGAA 279
 Db 173 -----CysGlnAsnValGlnSerGlnCysCysArgGlnLeuAl 185
 QY 280 CGAGTTTGAACAACCAAGGTGCATGTGCGAGGCATGTCACAGAGATTCGAGAGACCA 339

Db 185 aGlnIleProGluGlnPheArg-----CysGlnAlaIleHisAsnValValGlu----- 201
 QY 340 GAGCGATAGTTTCAGGGGAGCAACAGAGCAACAG-----TTCAA 381
 Db 202 -SerIleArgGlnGlnHisGlnProGlnGlnValGlnLeuGluGlyLeuAr 221
 QY 382 GAGGAGCTCAGAACTTGCCTCAACAGTGGCGCTTAGGGCCACACACAGCGTTGC 436
 Db 221 gMetSerLeuHisThrLeuProSerMetCysLysIleTyrIleProValGlnCys 239

RESULT 12

US-10-508-263-32
 ; Sequence 32, Application US/10508263
 ; Publication No. US20050260754A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BASF Plant Science GmbH
 ; TITLE OF INVENTION: Constructs and methods for regulating gene expression
 ; FILE REFERENCE: 53262-20085.00
 ; CURRENT APPLICATION NUMBER: US/10/508,263
 ; CURRENT FILING DATE: 2004-09-20
 ; NUMBER OF SEQ ID NOS: 126
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 32
 ; LENGTH: 178
 ; TYPE: PRT
 ; ORGANISM: Brassica napus

US-10-508-263-32

Alignment Scores:
 Pred. No.: 0.0158 Length: 178
 Score: 106.50 Matches: 40
 Percent Similarity: 41.3% Conservative: 31
 Best Local Similarity: 23.3% Mismatches: 56
 Query Match: 8.3% Indels: 45
 DB: 9 Gaps: 8

US-10-728-323-2 (1-1717) x US-10-508-263-32 (1-178)

QY 8 ATACTAGTAGCCCTCGCCCTTTCTCTCGTCCGCCCAGCATCTCGG---AGGCAGCAG 64
 Db 8 ValSerAlaThrLeuAlaPhePheLeuLeuThrAsnAlaSerIleTyrArgThrVal 27
 QY 65 TGGGAACCTCAAGGAGAC-----AGAAGATGC 91
 Db 28 valGluPheAspGluAspAlaThrAspSerAlaGlyProPheArgIleProLysCys 47
 QY 92 CAGAGCCAGCTCGAGAGGCG---AACCTGAGGCCCTCGCAGCAACATCTCATGCAGAAG 148
 Db 48 ArgLysGluPheGlnGlnAlaGlnHisLeuArgAlaCysGlnGlnTrpLeuHisLysGln 67
 QY 149 ATCCAACGTGAC-----GAGGAT 166
 Db 68 AlaMetGlnSerGlyGlyProSerTyrThrLeuAspGlyGluPheAspPheGluAsp 87
 QY 167 TCATATGAACGGGACCGGTACAGCCCTAGTCAGAGTCCGTACAGCCCTAGTCCATATGAT 226
 Db 88 AspMetGlu---AsnProGlnGlyProGlnGlnArgPro----- 99
 QY 227 CGGAGAGGGCGCTGGATCTCTCAGACCAAGAGAGGTGTTCGAATGAGCTGAACGAGTTT 286
 Db 100 -----ProLeuLeuGlnGlnCysCysAsnGluLeuHisGlnGlu 112
 QY 287 GAGAACCAACAAAGGTGCATGTCGAGGCATTCACACAGATCATGGAGACACGAGCGAT 346
 Db 113 Glu-----ProLeuCysValCysProThrLeuLysGlyAlaSerLysAlaValLysGln 130
 QY 347 AGGTTCGAGGGAGGCAACAGGAGCAACAGTTCAGAGGGAGCTCAGGAACCTTGCCTCAA 406
 Db 131 GlnIleGln---GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnMetValSerArgIleTyr 149
 QY 407 CAGTTCGGCCCTTAGGGCACACACAGCGTTGCGACTTG 442

Db 150 GlnThrAlaThrHisLeuProLysValCysAsnIle 161

RESULT 13

US-10-508-263-4
; Sequence 4, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-508-263-4

Alignment Scores:
Pred. No.: 0.0195 Length: 164
Score: 105.50 Matches: 40
Percent Similarity: 44.3% Conservative: 30
Best Local Similarity: 25.3% Mismatches: 59
Query Match: 8.2% Indels: 29
DB: 9 Gaps: 9

US-10-728-323-2 (1-717) x US-10-508-263-4 (1-164)

```

QY 8 ATACTAGTACCCCTCGCCCTTTCTCTCGCTGCGCCGATCTCGC---AGGCAGCAG 64
   ::::::::::::::::::::|||
Db 8 ValCysAlaThrLeuAlaLeuCysPheLeuLeuThrAsnAlaSerIleTyrArgThrVal 27
   ::::::::::::::::::::|||

QY 65 TGGGAACCTCAGGAGAC-----AGAGATGCCAGC 97
   |||:::|||||
Db 28 ValGluPheGluAspAlaSerAsnProValGlyProArgGlnArgCysGlnLys 47
   |||:::|||||

QY 98 CAGCTCAGAGGCGC---AACCTGAGCGCTCGGAGCAACATCTCATGCAAGATCCAA 154
   ::::::::::::::::::::|||
Db 48 GluPheGlnSerGlnHisLeuAlaCysGlnArgTrpMetSerLysGlnMetArg 67
   ::::::::::::::::::::|||

QY 155 CQTGACGAGATTATGAAACGGGACCCGTACAGCCCTAGTCAGGATCGGTACAGCCCT 214
   ::::::::::::::::::::|||
Db 68 GlnGlyArgGlyGly-----GlyProSerLeuAsp-----77

QY 215 AGTCCATATGATCGGAGAGCGCTGGATCTCT---CAGCACCAGAGAGGTGTTCAT 271
   ::::::::::::::::::::|||
Db 78 AspGluPheAspPheGluGlyProGlnGlnGlyTyrGlnLeuLeuGlnCysCysAsn 97
   ::::::::::::::::::::|||

QY 272 GAGCTCAACAGATTTCAGAACCAACCAAGGTGATGTGCGAGCATTTGCAACAGATCATG 331
   |||:::|||||
Db 98 GluLeuArgGlnGlu-----ProValCysValCysProThrLeuLysGlnAlaAla 115

QY 332 GAGAACACAGCATAGTTGTCAGGGAGGCA-----CAGGACCAACAGATTCAAG 382
   |||:::|||||
Db 116 ArgAlaValSer-----LeuGlnGlyGlnHisGlyProPheGlnSerArgLysIleTyr 133

QY 383 AGGAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGGGCACCACAGGTTGC 436
   ::::::::::::::::::::|||
Db 134 GlnSerAlaLysTyrLeuProAsnIleCysLysIleGlnGlnValGlyGluCys 151

```

RESULT 14

US-10-508-263-38
; Sequence 38, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Brassica napus cv. Topas
US-10-508-263-38

Alignment Scores:
Pred. No.: 0.0199 Length: 180
Score: 105.50 Matches: 44
Percent Similarity: 39.7% Conservative: 27
Best Local Similarity: 24.6% Mismatches: 49
Query Match: 8.2% Indels: 59
DB: 9 Gaps: 11

US-10-728-323-2 (1-717) x US-10-508-263-38 (1-180)

```

QY 8 ATACTAGTACCCCTCGCCCTTTCTCTCGCTGCGCCGATCTCGC---AGGCAGCAG 64
   ::::::::::::::::::::|||
Db 8 ValSerAlaThrLeuAlaPhePheLeuLeuThrAsnAlaSerIleTyrArgThrIle 27
   ::::::::::::::::::::|||

QY 65 TGGGAACCTCAGGAGAC-----AGAGATGC 91
   |||:::|||||
Db 28 ValGluValAspGluAspAlaThrAsnProAlaGlyProPheArgIleProLysCys 47
   |||:::|||||

QY 92 CAGAGCCAGCTCGAGAGGCGC---AACCTGAGCGCTCGGAGCAACATCTCATGCAAG 148
   ::::::::::::::::::::|||
Db 48 ArgLysGluPheGlnGlnAlaGlnHisLeuLysAlaCysGlnGlnTrpLeuHisLysGln 67
   ::::::::::::::::::::|||

QY 149 ATCCAACTGAC-----GAGGAT 166
   ::::::::::::::::::::|||
Db 68 AlaMetGlnSerGlySerGlyProSerTrpThrLeuAspGlyGluPheAspPheGluArg 87
   ::::::::::::::::::::|||

QY 167 TCATATGAGCGGACCCGTACAGCCCTAGTCAGGATCGGTACAGCCCTAGTCATATGAT 226
   |||:::|||||
Db 88 AspMetGlu---AsnProGlnGlyProGlnGlnArgPro-----99

QY 227 CGGAGAGCGCTGGATCTCTCAGCACCACCAAGAGAGGTGTTCATGAGTGAACGAGTTT 286
   ::::::::::::::::::::|||
Db 100 -----ProLeuLeuGlnGlnCysCysAsnGlnLeuHisGlnGlu 112

QY 287 GAGACAAACCAAGTGCATGTGCGAGGCAATTG-----CAA 322
   |||:::|||||
Db 113 Glu-----ProLeuCysValCysProThrLeuLysGlyAlaSerLysAlaValLysGln 130

QY 323 CAGATCATGGAACACAGAGCGATAGTTGCGAGGAGGACACAGGACCAAGTTCAAG 382
   |||:::|||||
Db 131 GlnValArgGlnGlnGlnGlyGln-----GlnGlyGlnGlnLeuGlnGlnValIleSer 148

QY 383 AGG-----GAGCTCAGGAACCTTGCTCAACAGTGGCGCTTAGGGCACCACAC 430
   |||:::|||||
Db 149 ArgIleTyrGlnThrAlaThrHisLeuProLysValCysAsnIle-----ProGln 165

```

RESULT 15

US-10-508-263-40
; Sequence 40, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Sinapis alba
US-10-508-263-40

Alignment Scores:
Pred. No.: 0.0269 Length: 144

```

Score: 104.00 Matches: 37
Percent Similarity: 42.0% Conservative: 23
Best Local Similarity: 25.9% Mismatches: 47
Query Match: 8.1% Indels: 36
DB: 9 Gaps: 7

US-10-728-323-2 (1-717) x US-10-508-263-40 (1-144)

QY 77 GGAGAGAGAGATGCCAGAGCCAGCTCGAGAGGGCG---AACCTGAGGGCCCTCGAGACAA 133
Db 6 GlyIleProLysCysArgLysGluPheGlnGlnAlaGlnHisLeuArgAlaCysGlnGln 25
QY 134 CATCTCATGCGAGATCCAACTG-----GACGAG 163
Db 26 TrpLeuHisLysGlnAlaMetGlnSerGlyProSerTrpThrLeuAspAspGlu 45
QY 164 GATTCTATATGAACGGGACCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATAT 223
Db 46 PheAspPheGluAspMetGluAsnPro---GlnGlyProGlnGlnArgProProLeu 64
QY 224 GATCGGAGAGCGCTGGATCCTCTCAGACCAAGAGAGAGGTGTGCAATGAGCTGAACGAG 283
Db 65 -----LeuGlnGlnCysAsnGluLeuHisGln 74
QY 284 TTTGAGAACACAAAGGTGTCATGTCGAGGCATTCACACAGATCATGGAG-----AAC 337
Db 75 GluGlu-----ProLeuCysValCysProThrLeuLysGlyAlaSerLysAlaValLys 92
QY 338 CAGAGCGATAGGTGTGCGAGGGAGGCAACAGGAGCAACAG-----376
Db 93 GlnGlnValArgGlnGlnLeuGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 112
QY 377 -----TTCAAGAGGGAGCTCAGGAATTCGCTCAACAGTCGGCCCTTAGGGCACCA 427
Db 113 IleSerArgIleTyrGlnThrAlaThrHisLeuProArgValCysAsnIleArgGlnVal 132
QY 428 CAGCGTTGC 436
Db 133 SerValCys 135

```

Search completed: May 16, 2006, 00:03:03
Job time : 15.3331 secs

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: May 15, 2006, 21:46:12 ; Search time 4.79902 Seconds
(without alignments)
4312.597 Million cell updates/sec

Title: US-10-728-323-2
Perfect score: 1281
Sequence: 1 gtcaccatactagtagcccc.....taaaagatcatgttttggct 717

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/abs/ABSWEB/spool/US10728323/runat_15052006_172134_22413/app_query.fasta_1
-DB=PIR -QFMT=fgstlan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US10728323 @CN1.1.92 @runat_15052006_172134_22413 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	278	21.7	153	A33090	conglutin delta pr
2	243	19.0	158	T05710	2S albumin precurs
3	187.5	14.6	80	B23617	conglutin delta-2
4	184.5	14.4	323	S38887	2S albumin - commo
5	173.5	13.5	141	T10257	2S albumin precurs
6	154	12.0	154	S14947	2S albumin - Brazil
7	151	11.8	295	S01062	2S seed storage pr
8	147	11.5	186	A29802	napin precursor (g
9	146.5	11.4	164	NWU01	2S albumin 1 precu
10	140.5	11.0	146	J14946	2S seed storage pr
11	139	10.9	155	JC5379	mabinlin II precu
12	139	10.9	258	RZCS	2S seed storage pr
13	132.5	10.3	165	T09252	seed storage prote
14	129	10.1	170	T08012	2S seed storage pr

15	128	10.0	172	2	S18871	2S-like storage pr
16	126.5	9.9	70	2	A59346	seed storage prote
17	125	9.8	162	2	T08013	2S seed storage pr
18	124.5	9.7	162	2	S49259	albumin 4 - easter
19	122.5	9.6	162	2	T08010	2S seed storage pr
20	121	9.4	100	2	S48180	mabinlin IV - Yun
21	121	9.4	174	2	PS0425	napin B43 precuso
22	119.5	9.3	180	2	S20225	napin (clones BNMN
23	118.5	9.3	106	2	S26636	napin n1b - rape
24	118.5	9.3	110	2	S20350	napin n1a - rape
25	118	9.2	104	2	S48176	mabinlin I-1 - Yun
26	117	9.1	104	2	S48178	mabinlin III - Yun
27	115	9.0	152	2	PS0427	napin AHI precuso
28	111.5	8.7	173	2	T08011	2S seed storage pr
29	108.5	8.5	178	2	S07828	napin B - rape
30	108.5	8.5	178	2	S25127	2S storage protein
31	107	8.4	139	2	T09850	albumin 2S storage
32	107	8.4	139	2	T09878	albumin 2S storage
33	106.5	8.3	178	1	NWRP2	napin 2 precursor
34	106.5	8.3	178	2	A25997	napin precursor (n
35	105.5	8.2	164	1	NWU03	2S albumin 3 precu
36	105.5	8.2	180	2	S10018	napin (clone BngNA
37	105.5	8.2	1099	2	A56155	tumor suppressor p
38	105.5	8.2	1909	2	A45592	liver stage antige
39	104.5	8.2	1263	2	T15496	hypothetical prote
40	104	8.1	145	2	S65479	allergen Sin a I (
41	103.5	8.1	166	1	NWU04	2S albumin 4 precu
42	103.5	8.1	178	2	S25130	2S storage protein
43	102.5	8.0	178	2	S25134	2S storage protein
44	101.5	7.9	178	2	S25137	2S storage protein
45	101	7.9	145	2	PC1246	Sin a I allergen 1

ALIGNMENTS

RESULT 1

A33090
conglutin delta precursor - narrow-leaved blue lupine
C:Species: Lupinus angustifolius (narrow-leaved blue lupine)
C>Date: 03-Aug-1990 #sequence_revision 03-Aug-1990 #text_change 09-Jul-2004
C:Accession: S12404; A33090
R:Gayler, K.R.; Kolivas, S.; Macfarlane, A.J.; Lilley, G.G.; Baldi, M.; Blagrove, R.J.; Plant Mol. Biol. 15, 879-893, 1990
A:Title: Biosynthesis, cDNA and amino acid sequences of a precursor of conglutin delta,
A:Reference number: S12404; MUID:91355912; PMID:2103479
A:Accession: S12404
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-153 <GAY>
A:Cross-references: UNIPROT:Q99235; UNIPARC:UPI000009D797; GB:X53523; NID:g19140; PIDN:
C:Superfamily: soybean 2S albumin
C:Keywords: seed
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-59/Product: conglutin small chain #status predicted <MAT1>
F:73-153/Product: conglutin large chain #status predicted <MAT2>

Alignment Scores:	Pred. No.:	Length:	Matches:
Score:	3 53e-18	278.00	62
Percent Similarity:	55.0%	Conservative:	26
Best Local Similarity:	38.8%	Mismatches:	46
Query Match:	21.7%	Indels:	26
DB:	2	Gaps:	5

US-10-728-323-2 (1-1717) x A33090 (1-153)

QY	2	CTCACCTACTAGTAGCCCTC---GCCCTTTTCTCTCGTCCGCCACCATCTGCAGG 58
		: : : : : : : : :
Db	4	LeuThrIleuIleAlaLeuValAlaAlaLeuValLeuValHisThrSerAlaPhe 23
		: : : : : : : :
QY	59	CAGCAGTGGGAACACTCCAGGACACAGACATCCAGACCTCAGAGGGGCGACCTTG 118
		: : : : : : : :
Db	24	Gln-----SerSerLysGlnSerCysLysArgGlnLeuGlnValAsnLeu 39
		: : : : : : : :

C;Accession: S38887
R;Thoyts, P.J.E.; Millichip, M.; Stobart, A.K.; Griffiths, W.T.; Napier, J.A.; Shewry, E.
submitted to the EMBL Data Library, November 1993
A;Reference number: S38887
A;Accession: S38887
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-323 <THO>
A;Cross-references: UNIPROT:Q39928; UNIPARC:UPI00000AC835; EMBL:X76101; NID:g429181; PID
C;Superfamily: gliadin

Alignment Scores:
Pred. No.: 2,52e-09 Length: 323
Score: 184.50 Matches: 46
Percent Similarity: 49.0% Conservative: 27
Best Local Similarity: 30.9% Mismatches: 43
Query Match: 14.4% Indels: 33
DB: 2 Gaps: 8

US-10-728-323-2 (1-717) x S38887 (1-323)

```
QY 83 AGAAGATGC---CAGAGCAGCTCGAGAGGCGAACCCTGAGGCCCTCGGAGCAACATCTC 139
      :::::||||| :::::||||| :::::||||| :::::||||| :::::|||||
Db 177 GlnGlnCysArgGlnSerGluIleGlnArgPro----ValSerGlnCysGlnArgGlyrVal 195
      :::::||||| :::::||||| :::::||||| :::::||||| :::::|||||
QY 140 ATGCAGAATCCAA-----CCTGACGAG 163
      :::::||||| :::::||||| :::::|||||
Db 196 GluGlnGlnIleGlnSerSerArgProTyrGlnGlnSerProTyrAspArgGlnGln 215
      :::::||||| :::::||||| :::::|||||
QY 164 GATTCTATGAAACGGACCGGTACAGCCCT-----AGTCAG 199
      :::::||||| :::::||||| :::::|||||
Db 216 SerProTyrAspArgGlnGlnSerProTyrGluGlnArgGlnGlyProTyrGluGln 235
      :::::||||| :::::||||| :::::|||||
QY 200 GATCCGTACAGCCCTAGTCCATATGATCGAGAGGCGCTGGATCCTCTCAGCACCAA--- 256
      :::::||||| :::::||||| :::::|||||
Db 236 ArgProTyrGluGlnArgProTyrGlnGlnArgGlyGlyArgGlnGlnGlnGlnGly 255
      :::::||||| :::::||||| :::::|||||
QY 257 ---GAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAACAAACCAAGGTGCATGTGCGAG 313
      :::::||||| :::::||||| :::::|||||
Db 256 LeuGlnGlnCysCysAsnGluLeuGlnAsnVal-----ArgArgGluCysGlnCysGlu 273
      :::::||||| :::::||||| :::::|||||
QY 314 GCATTCGAACAGATC-----ATGGAGAACAGACGATAGTTGCGAG 355
      :::::||||| :::::||||| :::::|||||
Db 274 AlaIleLysGluValGlyGlnArgMetArgGlnGlnGlnGlnArgArgGlnTyr 293
      :::::||||| :::::||||| :::::|||||
QY 356 GGGAGGCAACAGAGCAACAGTTCAAGAGGAGCTCAGGAACCTGCTCAACAGTCCGCG 415
      :::::||||| :::::||||| :::::|||||
Db 294 GlyGlyGlnGlnThrThrValGluArgGileLeuGluAsnLeuProAsnGlnCysAsp 313
      :::::||||| :::::||||| :::::|||||
QY 416 CTTAGGCGCACAGCGTTGCGACTTG 442
      :::::||||| :::::|||||
Db 314 LeuAspVal---GlnGlnCysAsnIle 321
      :::::||||| :::::|||||
```

RESULT 5

T10257
2S albumin precursor - cucurbit
N;Alternate names: prepro2S albumin
C;Species: Cucurbita sp. (cucurbit)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: T10257; S19323
R;Hara-Nishimura, I.; Takeuchi, Y.; Inoue, K.; Nishimura, M.
Plant J. 4, 793-800, 1993
A;Title: Vesicle transport and processing of the precursor to 2S albumin in pumpkin.
A;Reference number: Z17000; MUID:94100993; PMID:8275099
A;Accession: T10257
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA; protein
A;Residues: 1-141 <HARI>
A;Cross-references: UNIPARC:UPI00000A1B27; EMBL:DI6560; NID:g459404; PIDN:BAA03993.1; PI
A;Experimental source: seed, storage deposition stage; cotyledon
A;Note: soluble seed protein
R;Hara-Nishimura, I.; Inoue, K.; Nishimura, M.
FEBS Lett. 294, 89-93, 1991

A;Title: A unique vacuolar processing enzyme responsible for conversion of several prop.
A;Reference number: S19323; MUID:92077151; PMID:1743299
A;Accession: S19323
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 26-45;65-84 <HAR2>
A;Cross-references: UNIPARC:UPI000017CE18; UNIPARC:UPI000017CE19
C;Keywords: seed; storage protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-35/Domain: propeptide #status experimental <PRO>
F;36-141/Product: 2S albumin #status experimental <MAT>

Alignment Scores:
Pred. No.: 2,82e-08 Length: 141
Score: 173.50 Matches: 52
Percent Similarity: 49.4% Conservative: 26
Best Local Similarity: 32.9% Mismatches: 47
Query Match: 13.5% Indels: 33
DB: 2 Gaps: 10

US-10-728-323-2 (1-717) x T10257 (1-141)

```
QY 2 CTCACACTACTAGTAGCCCTC---GCCCTTTTCTCTCGCTGCCACCATCTGCGAGG 58
      :::::||||| :::::||||| :::::||||| :::::|||||
Db 4 LeuThrSerIleAlaLeuPheAlaValLeuLeuValAlaAspAlaTyrAlaTyr 23
      :::::||||| :::::||||| :::::|||||
QY 59 CAG-----CAGTGGGAACCTCCAAGGA---CACAGAACATGC 91
      :::::||||| :::::||||| :::::|||||
Db 24 ArgThrThrIleThrValGluValGluGluAsnArgGlnGlyArgGluGluArgCys 43
      :::::||||| :::::||||| :::::|||||
QY 92 CAGAGCCAGCTCGAGAGGCGAACCCTGAGGCCCTCGGACCAACATCTCTCATGCAAGATC 151
      :::::||||| :::::||||| :::::|||||
Db 44 ArgGlnMetSerAlaArgGluGluLeuArgSerCysGluGlnTyrLeuArgGlnGln--- 62
      :::::||||| :::::||||| :::::|||||
QY 152 CAACGTGACGAGGATTCATGAACGGGACCCGTCAGCCCTAGT-----CAGGATCCG 205
      :::::||||| :::::|||||
Db 63 -----SerArgAspValLeuGlnMetArgGlyIleGluAsnPro 75
      :::::||||| :::::|||||
QY 206 TACAGCCCTAGTCCATATGATCGAGAGCGCTGGATCTCTCAGACCAAGAGAGGTGT 265
      :::::||||| :::::|||||
Db 76 Trp-----ArgArgGluGlyGlySerPhe-----AspGluCys 86
      :::::||||| :::::|||||
QY 266 TGCATATGAGCTGAACGAGTTTGAGAACCAACAAAGGTGCATGTGCGAGCATTTGCAACAG 325
      :::::||||| :::::|||||
Db 87 CysArgGluLeuLysAsnValAsp-----GluGluCysArgCysAspMetLeuGluGlu 104
      :::::||||| :::::|||||
QY 326 ATCATGGAGAACCAAGCAGCATAGTTGACGGGAGGCGAACAGAGCAACAGTTCAAGAGG 395
      :::::||||| :::::|||||
Db 105 IleAlaArgGluGluGlnArgGlnAlaArgGly---GlnGluGlyArgGlnMetLeuGln 123
      :::::||||| :::::|||||
QY 386 GAGCTCAGGAACCTGCCTCAACAGGTGCGGCCCTTAGGGCCACACAGCGTTTGGCAG 439
      :::::||||| :::::|||||
Db 124 LysAlaArgAsnLeuProSerMetCysGlyIleArg---ProGlnArgCysAsp 140
      :::::||||| :::::|||||
```

RESULT 6

S14947
2S albumin - Brazil nut
C;Species: Bertholletia excelsa (Brazil nut)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 31-Dec-2004
C;Accession: S14947
R;Gander, E.S.; Holmstroem, K.O.; de Paiva, G.R.; de Castro, L.A.B.; Carneiro, M.; Gros
Plant Mol. Biol. 16, 437-448, 1991
A;Title: Isolation, characterization and expression of a gene coding for a 2S albumin f.
A;Reference number: S14946; MUID:91370890; PMID:1840683
A;Accession: S14947
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-154 <GAN>
A;Cross-references: UNIPARC:UPI000016DCCA; EMBL:X54491; NID:g17712; PIDN:CAA38363.1; PI
C;Genetics: 67/3
A;Introns:
C;Superfamily: Alpha amylase inhibitor

Alignment Scores:
Pred. No.: 1.99e-06 Length: 154
Score: 154.00 Matches: 45
Percent Similarity: 45.3% Conservatives: 28
Best Local Similarity: 28.0% Mismatches: 52
Query Match: 12.0% Indels: 36
DB: 2 Gaps: 10

US-10-728-323-2 (1-717) x S14947 (1-154)

QY 8 ATACTAGTAGCCTCGCCCTTTTCCTCTCGTCCCGCAGCATCTGCG----- 55
DB 7 ValAlaAlaLeuLeuAlaLeuLeuValLeuGlyGlnAlaThrAlaPheArgThrThr 26
QY 56 -----AGGCAGCAGTGGGAACCTCCAAGGA-----GACAGAAGA 88
DB 27 ValThrThrLeuGluGluGluGluGluAsnProArgGlyArgSerGluGln 46
QY 89 TGCAGAGCCAGCTCGAGG-----GCGAACCCTGAGCCCTGCGAGCAACATCTCATGCGAG 145
DB 47 CysArgGluGlnMetGluArgGlnGlnGlnLeuAsnHisCysArgMetTyrLeuArgGln 66
QY 146 AAGATCCAACTGACGAGGATTATATGAACGGGACCCGTACAGCCCTAGTCAGGATCCG 205
DB 67 GlnMet-----GluGluSerProTyrGln----- 74
QY 206 TACAGCCCTAGTCCATATGATCGGAGGCGCTGGATCTCTCAGCACCAAGAGAGGTGT 265
DB 75 ---AsnProArgProLeuArgA-g-----GlyGluGluProHisLeuAspGluCys 90
QY 266 TGAATGAGCTGACGAGTTTGAGAACCAACCAAGGTGATGTCGAGGCAATGTCACACAG 325
DB 91 CysGluGluLeuGluArgMetAsp-----GluMetCysArgCysGluGlyLeuArgMet 108
QY 326 ATCATGAGAACACAGAGCGAT-----AGGTTCAGGGGAGGCAACAGGAGCAACAGTTTC 379
DB 109 MetLeuArgArgGlnArgGluGluMetGluLeuGlnGly---GluGlnMetGlnArgIle 127
QY 380 AAGAGGGAGCTCAGGAATTCCTCAACAGTGGCGCTTGGGCAACACAGCTTGGCAG 439
DB 128 MetArgLysAlaGluAsnLeuSerArgCysAsnLeu---SerProGlnArgCysPro 146
QY 440 TTG 442
DB 147 Met 147

RESULT 7
S01062
2S seed storage protein precursor (clone HaGS) - common sunflower
N;Alternate names: 2S albumin storage protein
C;Species: Helianthus annuus (common sunflower)
C;Date: 31-Dec-1998 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C;Accession: S01062
R;Allen, R.D.; Cohen, E.A.; Vonder Haar, R.A.; Adams, C.A.; Ma, D.P.; Nessler, C.L.; Tho
Mol. Gen. Genet. 210, 211-218, 1987
A;Title: Sequence and expression of a gene encoding an albumin storage protein in sunfl
A;Reference number: S01062; MUID:88142538; PMID:2830455
A;Accession: S01062
A;Molecule type: DNA
A;Residues: 1-295 <ALL>
A;Cross-references: UNIPROT:P15461; UNIPARC:UPI0000124EBB; EMBL:X06410; NID:g18799; PIDN
A;Note: part of this sequence, including the amino end of the mature protein, was confir
C;Genetics:
F;1-20/Domain: signal sequence #status predicted <SIG>
F;162-295/Product: 2S seed storage protein #status experimental <MAT>

Alignment Scores:
Pred. No.: 3.77e-06 Length: 295
Score: 151.00 Matches: 47
Percent Similarity: 40.4% Conservatives: 35
Best Local Similarity: 23.2% Mismatches: 65

Query Match: 11.8% Indels: 56
DB: 2 Gaps: 8

US-10-728-323-2 (1-717) x S01062 (1-295)

QY 8 ATACTAGTAGCCTCGCCCTTTTCCTCTCGTCCCGCAGCATCTGCG----- 55
DB 5 IleValLeuAlaLeuAlaPheAlaLeuValAlaPheAlaThrAlaHisThrIle 24
QY 56 -----AGGCAGCAGTGGGAACCTCCAAGGAGACAGAGATGCCAGAGCCAG 100
DB 25 IleThrThrThrIleGluAspGluAsnProIleSerGlyGlnArgGlnValSerGlnArg 44
QY 101 CTCGAGAGGCGCAACCTGAGGCCCTCGGAGCAACATCTCTCATG-----CAGAAGATCCAA 154
DB 45 IleGlnGlyGlnArgLeuAsnGlnCysArgMetPheLeuGlnGlnGlyGlnAsnIlePro 64
QY 155 CGTGCAGGAGATTTCATATGAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCT 214
DB 65 ArgGluPheAspAsn----- 69
QY 215 AGTCCATATGATCGGAGAGGCGCTGGATCCTCTCAGCACCAAGAGAGGTGTTCGAATGAG 274
DB 70 -----ProGlnMetGlyArgGlnGlnGlnGlnLeuGlnCysCysGlnGlu 86
QY 275 CTGACAGCTTTGAGAACCAACAAAGGTGCATGTGCGAGCATTCGACAGATCATCGGAG 334
DB 87 LeuGlnAsnIleGly-----GlnCysGlnCysGluAlaValLysGlnValPheArg 104
QY 335 AACACAGGCGATAGTTGCGAG-----GGGAGG 361
DB 105 GluAlaGlnGlnValGlnGlnGlnGlnGlnGlnGlnValPropheArgGlySer 124
QY 362 CRACAGGAGCAACAGTTCAAGAGGAGGAGCTCAGGAACCTGCCTCAACAGTGCAGCCCTTAGG 421
DB 125 GlnGlnThrGlnGlnLeuLysGlnLysAlaGlnIleLeuProAsnValCysAsnLeuGln 144
QY 422 GCACACAGCGTTGCCACTTG-----GAGCTCGAA 451
DB 145 Ser---ArgArgCysGluIleGlyThrIleThrThrThrValThrGluSerAsnIleAsp 163
QY 452 AGTGGCGGAGACAGACAGATATAACACCTATCTCAAAAAAGAGAAAGAAAGAAAGAA 511
DB 164 IlePropheArgAspArg-----PropheGlyThrGlySerGlnGlnCysArgGlu 180
QY 512 AATAGCTTA 520
DB 181 ThrGluIle 183

RESULT 8
A29802
napin precursor (gNa) - rape
N;Alternate names: 1.7 S seed storage protein
C;Species: Brassica napus (rape)
C;Date: 20-Jun-1989 #sequence_revision 20-Jun-1989 #text_change 31-Dec-2004
C;Accession: A29802
R;Scofield, S.R.; Crouch, M.L.
J. Biol. Chem. 262, 12202-12208, 1987
A;Title: Nucleotide sequence of a member of the napin storage protein family from Brass
A;Reference number: A29802; MUID:87308225; PMID:3040733
A;Accession: A29802
A;Molecule type: DNA
A;Residues: 1-186 <SCO>
A;Cross-references: UNIPROT:P09893; UNIPARC:UPI0000124EBE; GB:J02782; NID:g167178; PIDN
C;Genetics:
C;Superfamily: Alpha amylase inhibitor

Alignment Scores:
Pred. No.: 9.11e-06 Length: 186
Score: 147.00 Matches: 47
Percent Similarity: 41.0% Conservatives: 28
Best Local Similarity: 25.7% Mismatches: 52

Query Match:	11.5%	Indels:	56
DB:	2	Gaps:	9
US-10728-323-2 (1-717) x A29802 (1-186)			
QY	8	ATACTAGTAGCCCTCGCCCTTTCTCCTCGCTGCCCCACGCATCTCGC	---AGGCAGCAG 64
DB	8	ValSerAlaThrLeuAlaLeuPheLeuLeuThrAsnAlaSerValTyArgThrVal	27
QY	65	TGGGAACCTCCAAGGAGAC	-----AGAAGATGC 91
DB	28	ValGluValAspGluAspAlaThrAsnProAlaGlyProPheArgIleProLysCys	47
QY	92	CAGAGCCAGCTCGAGGGCG	---AACTGAGGCCCTCGCAGCAACAATCTTCATGCAGAAG 148
DB	48	ArgLysGluPheGlnGlnAlaGlnHisLeuArgAlaCysGlnInTrpLeuHisLysGln	67
QY	149	ATCCAACGTGAC	----- 160
DB	68	AlaMetGlnProGlyGlySerGlyProSerTrpThrLeuAspGlyGluPheAspPhe	87
QY	161	GAGGATTCAATGAAACGGGACCCGTACAGCCCTAGTCAGGATCGTACAGCCCTAGTCCA	220
DB	88	GluAspAspValGluAsnGlnGlnGlyProGlnGlnArgPro	-----PropProPro 105
QY	221	TATGATCGGAGAGCGCTGGATCCTCTCAGCACCAAGAGAGAGGTGTTGCAATGAGCTGAAC	280
DB	106	-----GlnGlnCysCysAsnGluLeuHis	113
QY	281	GAGTTTGAGAACCAACAAAGTGTCATGTGGAGCATTG	----- 319
DB	114	GlnGluGlu-----ProLeuCysValCysProThrLeuLysGlyAlaSerLysAlaVal	131
QY	320	---CAACAGATCATGGAGAACCCAGAGCATAGTTGACGGGGAGGCAACAGGAGCAACAG	376
DB	132	ArgGlnGlnValArgGlnGlnGlnGlyGlnGlnMetGlnGlnMetGlnGlnVal	151
QY	377	TTCAGAGGGAGCTCAGG	-----AACTTGCTCAACAGTGCAGCGCTTATGGGCACCA 427
DB	152	IleSerArgValTyGlnThrAlaThrHisLeuProArgValCysAsnIleArgGlnVal	171
QY	428	CAGCGTTGC 436	
DB	172	SerIleCys 174	
RESULT 9			
NWU1			
2S albumin 1 precursor - Arabidopsis thaliana			
N:Alternate names: seed storage protein AT2S1			
C:Species: Arabidopsis thaliana (mouse-ear cress)			
C:Date: 31-Mar-1992 #sequence revision 31-Mar-1992 #text_change 31-Dec-2004			
C:Accession: JA0161; PS0282; S34676; T06044			
R:Krebbes, E.; Herdies, L.; de Clercq, A.; Seurinck, J.; Leemans, J.; Van Damme, J.; Se			
Plant Physiol. 87, 859-866, 1988			
A:Title: Determination of the processing sites of an Arabidopsis 2S albumin and character			
A:Reference number: JA0161			
A:Accession: JA0161			
A:Molecule type: DNA			
A:Residues: 1-164 <KRE>			
A:Cross-references: UNIPROT:P15457; UNIPARC:UPI0000000ABC; GB:M22032; NID:g166609; PIDN:			
A:Accession: PS0282			
A:Molecule type: protein			
A:Residues: 38-73; 84-162 <KR2>			
A:Cross-references: UNIPARC:UPI0000173293; UNIPARC:UPI0000173294			
R:Conceicao, A.D.S.; Krebbes, E.			
Submitted to the EMBL Data Library, July 1993			
A:Description: Tentative title: a cotyledon regulatory region is responsible for the dif			
A:Reference number: S34674			
A:Accession: S34676			
A:Molecule type: DNA			
A:Residues: 1-164 <CON>			
A:Cross-references: UNIPARC:UPI0000000ABC; EMBL:Z24745; NID:g395203; PIDN:CAA80870.1; PI			
R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Ba			

submitted to the Protein Sequence Database, March 1999

A:Reference number: Z15484
A:Accession: T06044
A:Molecule type: DNA
A:Residues: 1-164 <BEV>
A:Cross-references: UNIPARC:UPI0000000ABC; EMBL:AL035680; GSPDB:GN00062; ATSP:T24A18.90
A:Experimental source: cultivar Columbia; BAC clone T24A18
C:Genetics:
A:Gene: T24A18.90
A:Map position: 4
C:Keywords: seed; storage protein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-164/Product: 2S albumin 1 proprotein #status predicted <AT2>
F:38-73/Product: 2S albumin 1 small chain #status experimental <SMC>
F:84-162/Product: 2S albumin 1 large chain #status experimental <LAC>

Alignment Scores:
Pred. No.: 1.02e-05 Length: 164
Score: 146.50 Matches: 46
Percent Similarity: 46.5% Conservative: 33
Best Local Similarity: 27.1% Mismatches: 52
Query Match: 11.4% Indels: 39
DB: 1 Gaps: 9

US-10-728-323-2 (1-717) x NMWU (1-164)

Qy	8	ATACTAGTAGGCGCTCGCCCTTTCTCTCTCTGCGCCACGATCTGCG---	AGGCAGCAG	64
Db	8	ValCysAlaLaLeuAlaLeuCysPheLeuThrAsnAlaSerIleTyrArgThrVal	27	
Qy	65	TGGGAATCTCAAGGAC---	-----AGAAGATGCCAG	94
Db	28	ValGluPheGluGluAspAlaThrAsnProIleGlyProLysMetArgLysCysArg	47	
Qy	95	AGCCAGCTCGAGAGG---GCCAACTGAGGCCCTCGGAGCAACATCTCATGCAGAGATC	151	
Db	48	LysGluPheGlnLysGluGlnHisLeuArgAlaCysGlnGlnLeuMetLeuGlnAla	67	
Qy	152	CAACGT-----GACGAGGATTCATATCAAGCGGACCGGTACAGCCCTAGTCAGGAT	202	
Db	68	ArgGlnGlyArgSerAspGluPheAspPheGluAspMetGluAsnProGln-----	85	
Qy	203	CCGTACAGCCCTAGTCCCATATGATCGGAGAGCGCTGGATCCTCTCAGACCAA-----	256	
Db	86	-----GlyGlnGlnGlnGlnLeu	93	
Qy	257	---GAGAGGTGTTGCATGACCTGAAACGAGTTTGAGAACACCAAGTGCATGTGCCAG	313	
Db	94	PheGlnGlnCysCysAsnGluLeuArgGlnGluGluProAsp-----CysValCysPro	111	
Qy	314	GCATTGCAACAGATCATGGAGAACACAGAGCGATAGTTTGACGGGAGGCAACAGGCAA	373	
Db	112	ThrLeuLysGlnAlaAla-----LysAlaValArgLeuGlnGlnHisGlnProMet	129	
Qy	374	CAGTTCAAGAGG-----GAGCTCAGGAACCTTGCTCAACAGTGGCGCTTGAAGCA	424	
Db	130	GlnValArgLysIleTyrGlnThrAlaLysHisLeuProAsnValCysAspIleProGln	149	
Qy	425	CCACAGCGTTCCGACTTGGACGTCGAAGT	454	
Db	150	ValAspValCysProPheAsnIleProSer	159	

RESULT 10
SI4946
2S seed storage protein large chain - Brazil nut
N:Alternate names: albumin 2S precursor
C:Species: Bertholletia excelsa (Brazil nut)
C:Date: 21-Nov-1993 #sequence revision 26
C:Accession: SI4946; SI4479; S06252; S21640; B25802
R:Gander, E.S.; Holmstrom, K.O.; de Paiva, G.R.; de Castro, L.A.B.; Carneiro, M.; Gross
Plant Mol. Biol. 16, 437-448, 1991
A:Title: Isolation, characterization and expression of a gene coding for a 2S albumin f
A:Reference number: SI4946; MUID:91370890; PMID:1840683

```

Db      63 -----GluSerProTyrGlnThrMet 69
QY    218 CCAATATGATCGAGAGCGCTGGATCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTG 277
      ||| ||||| |||||
Db     70 Pro-----ArgArg-----GlyMetGluProHisMetSerGluCysCysGluGlnLeu 85
      ||| ||||| |||||
QY    278 AACGAGTTTGAGAACAACAAAGGTGCATGTGGCAGGCATTGCCAA---CAGATCATGGAG 334
      ::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db     86 GluGlyMetAsp-----GluSerCysArgCysGluGlyLeuArgMetMetMetArg 103
      ::: :::: :::: :::: :::: :::: :::: :::: ::::
QY    335 AACCAGAGCAGTAGTGTTCAGGGGAGG---CAAACAGAGCAACAGTTCACAGAGGAGCTC 391
      ||| :::: :::: :::: :::: :::: :::: ::::
Db    104 MetGlnGlnGluMetGlnProArgGlyGluGlnMetArgMetMetArgLeuAla 123
      ||| :::: :::: :::: :::: :::: ::::
QY    392 AGGAACCTTGCTCAACAGATGCGCCCTTAGGCGCACACAGGCTTCGCACTTG 442
      ||| :::: :::: :::: :::: :::: ::::
Db    124 GluAsnIleProSerArgCysAsnLeu---SerPrometArgCysPromet 139

RESULT 11
mabinlin II precursor - Yunnan caper
C:Species: Capparis maseikai (Yunnan caper)
C>Date: 02-Jun-1997 #status experimental <MAT>
C:Accession: JCS379; PC4316
R:Nirasawa, S.; Masuda, Y.; Nakaya, K.; Kurihara, Y.
Gene 181, 225-227, 1996
A:Title: Cloning and sequencing of a cDNA encoding a heat-stable sweet protein.
A:Reference number: JCS379; MUID:97128796; PMID:8973336
A:Accession: JCS379
A:Molecule type: mRNA
A:Residues: 1-155 <NIR1>
A:Cross-references: UNIPROT:P30233; UNIPARC:UPI000016DD17; DBJ:D83997; NID:g1g1
A:Accession: PC4316
A:Molecule type: protein
A:Residues: 36-41;149-154 <NIR2>
A:Cross-references: UNIPARC:UPI00001763F3; UNIPARC:UPI00001763F4
A:Experimental source: seed
C:Keywords: sweet taste
C:Keywords: signal sequence #status predicted <SIG>
F:1-20/Domain: amino-terminal propetide #status predicted <PRO>
F:21-35/Domain: amino-terminal propetide #status predicted <PRO>
F:36-68,83-154/Product: mabinlin II #status experimental <MAT>
F:36-68/Domain: mabinlin II A chain #status experimental <CHA>
F:69-83/Domain: mabinlin linker peptide #status predicted <LNK>
F:83-154/Domain: mabinlin II B chain #status experimental <CHB>

Alignment Scores:
Pred. No.:          5.24e-05           Length:         155
Score:             139.00              Matches:         45
Percent Similarity: 45.5%              Conservative:    26
Best Local Similarity: 28.8%            Mismatches:     47
Query Match:       10.9%                Indels:         38
DB:                                                         Gaps:          9

US-10-728-323-2 (1-717) x JCS379 (1-155)

QY    8 ATACTAGTACCCCTCGCCCTTTCTCCTCTCGCTGCCACGCATCTCGCAGG----- 58
      ::: ||||||| ||||| :::: ||||| ||||| |||||
Db   7 LeuPheAlaThrLeuAlaLeuPheValLeuAlaSerIleGlnThrVal 26
      ||| ||||| ||||| ||||| ||||| ||||| |||||
QY   59 -----CAGCAGTGGGAACCTCCAAGGAGACAGAATGC 91
      ||||| ||||| ||||| ||||| ||||| |||||
Db   27 IleGluValAspGluGluAspAsnGlnLeuTrp-----ArgCys 40
      ||||| ||||| ||||| ||||| ||||| |||||
QY   92 CAGAGCCAG---CTCGAGAGGGCGAACCTGTAGGCCCTCGGAGCAACATCTCATGCAGAG 148
      ||||| ||||| ||||| ||||| ||||| |||||
Db   41 GlnArgGlnPheLeuGlnHisGlnArgLeuAlaCysGlnArgPheIleHisArg 60
      ||||| ||||| ||||| ||||| ||||| |||||
QY  149 ATCCAA-----CGTGACGAGGATTCATATCAACCGGACCCGCTACAGCCCTAGTCAGGAT 202
      ||||| ||||| ||||| ||||| ||||| |||||
Db   61 AlaGlnPheGlyGlnProAspGluLeuGluaspGluValGluAspAsnAsp 80
      ||||| ||||| ||||| ||||| ||||| |||||
QY  203 CCGTACAGCCCTAGTCCATATGATCGGAGAGCGCGCTGTGATCCTCTCAGCACCAAGAGG 262
      ||||| ||||| ||||| ||||| ||||| |||||

```



```

Query Match: 10.3% Indels: 65
DB: 2 Gaps: 10

US-10-728-323-2 (1-717) x T09252 (1-165)

QY 8 ATACTAGTACCTCGCCCTTTCTCTCGCT-----GCC 43
   ::: ::: |||||
Db 17 ValSerLeuGlyMetAlaLeuLeuLeuLeuLeuHisTrpGlyThrArgThrValAspAla 36
   |||
QY 44 CAGCATCTCGAGGCGAGTGGAACTCCAAGGAGACAGAAGA---TGCCAGAGCCAG 100
   |||
Db 37 HisGluAspGlyLeuTyGlyGluGluValGlnGlnArgSerCysGluGlnGln 56
   |||
QY 101 CTCGAGAGGCGGAACCTGAGCCCTCGGACCAACTCTCATGCAAGATCCCAACGTGAC 160
   |||
Db 57 -----ArgLeuSerSerCysArgGluTyrlu-----GluArgPro 68
   |||
QY 161 GAGGATTATGAACGGGACCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCA 220
   |||
Db 69 ArgAsp----- 70
QY 221 TATGATCGGAGAGCGCTGGATCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAAC 280
   |||
Db 71 -----GlnProSerGluArgCysGluGluLeuGln 81
QY 281 GAGTTTCAGAACCAACAGGTGCTGCGAGCATTCGAACAGATCATGGAAGAAC--- 337
   |||
Db 82 ArgMet-----SerProGlnCysArgCysGlnAlaIleGlnArgThrLeuGluAspVal 99
   |||
QY 338 -----CAGAGCGAT-----AGG 349
   |||
Db 100 PheMetAspSerAspSerGlnAspGlyAlaProLeuAsnGlnArgArgGlnArgArg 119
QY 350 TTCAGGGGAGG---CAACAGGACCAACAGTTCAGAGGAGGCTCAGGAACCTTCCTCAA 406
   |||
Db 120 GlyGlnGlyArgGlyMetGluGluGluValValArgAlaGluGluLeuProAsn 139
QY 407 CAGTGGCGCTTAGG---GCACCACAGCGTTGGACTTGCAGCTCGNAAGT 454
   |||
Db 140 ThrCysAsnValArgGlnSerProArgArgCysAspLeuGlnArgHisSer 156

RESULT 14
T08012
2S seed storage protein precursor - Douglas fir
C;Species: Pseudotsuga menziesii (Douglas fir)
C;Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 31-Dec-2004
C;Accession: T08012
R;Chatthai, M.; Mistra, S.
Planta 206, 138-145, 1998
A;Title: Structure and expression of embryogenesis-specific cDNAs encoding 2S seed storage
A;Reference number: Z16286; MUID:98381325; PMID:9715536
A;Accession: T08012
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-170 <CHA>
A;Cross-references: UNIPROT:O64931; UNIPARC:UPI000000AA754; EMBL:AF029972; NID:G2988481;
C;Genetics:
A;Gene: 2S3
C;Superfamily: 2S albumin-like seed storage protein
C;Keywords: seed; storage protein

Alignment Scores:
Pred. No.: 0.000464 Length: 170
Score: 129.00 Matches: 43
Percent Similarity: 36.1% Conservative: 22
Best Local Similarity: 23.9% Mismatches: 49
Query Match: 10.1% Indels: 66
DB: 2 Gaps: 8

US-10-728-323-2 (1-717) x T08012 (1-170)

QY 8 ATACTAGTACCTCGCCCTTTCTCTCGCT-----GCC 43
   ::: ::: |||||
Db 17 ValSerLeuGlyMetAlaLeuLeuLeuLeuLeuHisTrpGlyThrArgThrValAspAla 36
   |||
QY 44 CAGCATCTCGAGGCGAGTGGAACTCCAAGGAGACAGAAGA---TGCCAGAGCCAG 100
   |||
Db 37 HisGluAspGlyLeuTyGlyGluGluValGlnGlnArgSerCysGluGlnGln 56
   |||
QY 101 CTCGAGAGGCGGAACCTGAGCCCTCGGACCAACTCTCATGCAAGATCCCAACGTGAC 160
   |||
Db 57 -----ArgLeuSerSerCysArgGluTyrlu-----GluArgPro 68
   |||
QY 161 GAGGATTATGAACGGGACCGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCA 220
   |||
Db 69 ArgAsp----- 70
QY 221 TATGATCGGAGAGCGCTGGATCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAAC 280
   |||
Db 71 -----GlnProSerGluArgCysGluGluLeuGln 81
QY 281 GAGTTTCAGAACCAACAGGTGCTGCGAGCATTCGAACAGATCATGGAAGAAC--- 337
   |||
Db 82 ArgMet-----SerProGlnCysArgCysGlnAlaIleGlnArgThrLeuGluAspVal 99
   |||
QY 338 -----CAGAGCGAT-----AGG 349
   |||
Db 100 PheMetAspSerAspSerGlnAspGlyAlaProLeuAsnGlnArgArgGlnArgArg 119
QY 350 TTCAGGGGAGG---CAACAGGACCAACAGTTCAGAGGAGGCTCAGGAACCTTCCTCAA 406
   |||
Db 120 GlyGlnGlyArgGlyMetGluGluGluValValArgAlaGluGluLeuProAsn 139
QY 407 CAGTGGCGCTTAGG---GCACCACAGCGTTGGACTTGCAGCTCGNAAGT 454
   |||
Db 140 ThrCysAsnValArgGlnSerProArgArgCysAspLeuGlnArgHisSer 156

RESULT 15
S18871
2S-like storage protein - white spruce
C;Species: Picea glauca (white spruce)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 31-Dec-2004
C;Accession: S18871
R;Newton, C.H.
submitted to the EMBL Data Library, November 1991
A;Reference number: S18871
A;Accession: S18871
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-172 <NEW>
A;Cross-references: UNIPROT:P26986; UNIPARC:UPI0000124EC0; EMBL:X63193; NID:G20495; PID
C;Superfamily: 2S albumin-like seed storage protein

Alignment Scores:
Pred. No.: 0.000576 Length: 172
Score: 128.00 Matches: 46
Percent Similarity: 37.4% Conservative: 21
Best Local Similarity: 25.7% Mismatches: 46
Query Match: 10.0% Indels: 66
DB: 2 Gaps: 9

US-10-728-323-2 (1-717) x S18871 (1-172)

QY 14 GTAGCCCTCGCCCTTTCTCTCTCTC-----GCTGCCACGCA 49
   ::: ::: |||||
Db 18 LeuSerValAlaLeuPheLeuLeuPheHisTrpGlyLeProSerValAspGlyHisGlu 37
   |||
QY 50 TCTCGAGGCGAGTGGGAATCTCAAGGAGACAGAAGA---TGCCAGAGCCAGCTCGAG 106
   |||
Db 38 AspAsnMetTyrluGluGluLeuGlnGlnArgSerCysAspProGlnArgAsp 57
   |||
QY 107 AGGCGCAACCTGAGGCCCTCGGACCAACTCTCATGCAAGATCCCAACGTGACGAGAT 166
   |||
Db 58 ProGlnArgLeuSerSerCysArgAspTyrlu----- 68
QY 167 TCATATGAACGGGACCGGTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCATATGAT 226
   |||

```

```

Db      69  -----GluArgArgGluGlnProSer----- 76
QY      227  CGGAGAGCGCTGGATCCTCTCAGCACCAAGAGAGCTGTTCGAATCAGCTGACGAGTTT 286
Db      77  -----GluArgCysCysGluGluLeuGlnArgMet 86
QY      287  GAGAACAAACAAAGGTGCATGTGCGAGGCATTGCAACAGATCATGGAG----- 334
Db      87  -----SerProGlnCysArgCysGlnAlaIleGlnMetLeuaspGlnSerLeuSer 104
QY      335  -----AACGAGCGGATAGG 349
Db      105  TyrAspSerPheMetAspSerAspSerGlnGluAspAlaProLeuAsnGlnArgArg 124
QY      350  TTG-----CAGGGAGGCAACAGAGCAACAGTTTCAAGAGGGAGCTCAGGAACCTG 400
Db      125  ArgArgGluGlyArgGluGluGluAlaMetGluArg---AlaAlaTyrLeu 143
QY      401  CCTCAACAGTGGCGCCCTTAGG---GCACCACAGCGTTGCGACTTGGACGTCGAAAGT 454
Db      144  ProAsnThrCysAsnValArgGluProProArgArgCysAspIleGlnArgHisSer 162

```

Search completed: May 15, 2006, 22:21:58
 Job time : 26.9951 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 21:23:21 ; Search time 32.6702 Seconds
(without alignments)

4645.181 Million cell updates/sec

Title: US-10-728-323-2

Perfect score: 1281

Sequence: 1 gctcaccatactagtagcc.....taaaagatcatgtttgtt 717

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DBV=xlp
-Q=/abes/ABESWEB_spool/US10728323/runat_15052006_172132_22377/app_query_fasta_1
-DB=UniProt -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abes02p
-USER=us10728323 @CCN 1.1 694 @runat_15052006_172132_22377 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	842	65.7	179	2 Q647H0	ARAHY Q647h0 arachis hyp
2	838	65.4	175	2 Q6PSU1	ARAHY Q6psu1 arachis hyp
3	819	63.9	156	2 Q941R0	ARAHY Q941r0 arachis hyp
4	815	63.6	172	2 Q8GV20	ARAHY Q8gv20 arachis hyp
5	815	63.6	187	2 Q6PSU2	ARAHY Q6psu2 arachis hyp
6	808	63.1	169	2 Q7Y1C0	ARAHY Q7y1c0 arachis hyp
7	792	61.8	166	2 Q84TU1	ARAHY Q84tu1 arachis hyp
8	454.5	35.5	145	2 Q647G9	ARAHY Q647g9 arachis hyp
9	443.5	34.6	144	2 Q8W251	ARAHY Q8w251 arachis hyp
10	443.5	34.6	145	2 Q51G71	ARAHY Q51g71 arachis hyp
11	401	31.3	124	2 Q511E9	ARAHY Q511e9 arachis hyp
12	388.5	30.3	129	2 Q9SQG5	ARAHY Q9sqg5 arachis hyp
13	300	23.4	160	2 Q9SQH1	ARAHY Q9sqh1 arachis hyp
14	299	23.3	158	2 Q647G8	ARAHY Q647g8 arachis hyp
15	278	21.7	153	2 Q99235	LUPAN Q99235 lupinus ang
16	243	19.0	158	1 2SS_SOYN	P19594 glycine max

17	243	19.0	158	2 Q53WV6	SOYBN Q53wv6 glycine max
18	241.5	18.9	155	2 Q9ZN24	SOYBN Q9zn24 glycine max
19	187.5	14.6	80	1 CGD2L	LUPAN CGD2l lupinus ang
20	184.5	14.4	323	2 Q3928	HELAN Q3928 helianthus
21	177.5	13.9	168	2 Q8LPD3	LINUS Q8lpd3 linum usita
22	173.5	13.5	141	1 2SS_CUGMA	Q3949 cucurbita m
23	167	13.0	160	2 Q71HN1	FICAW Q71hn1 ficus awkeo
24	164.5	12.8	148	2 Q9XHP1	SESIN Q9xhp1 sesamum ind
25	159.5	12.5	143	2 Q7YK62	SEGIN Q7yk62 sesamum ind
26	159	12.4	161	2 Q7Y1C2	JUGNI Q7y1c2 juglans nig
27	153.5	12.0	139	2 Q93198	GR0SI Q93198 juglans reg
28	153	11.9	153	2 Q9AUD1	SESIN Q9aud1 sesamum ind
29	151	11.8	295	1 2SS5_HELAN	P15461 helianthus
30	149.5	11.7	167	2 Q84JW2	VITVI Q84jw2 vitis vinif
31	148.5	11.6	140	2 Q8L694	MOMCH Q8l694 momordica c
32	148	11.6	285	2 Q8GUD8	HELAN Q8gud8 helianthus
33	147	11.5	186	1 2SSE_BRANA	P09893 brassica na
34	146.5	11.4	164	1 2SSI_ARATH	P15457 arabidopsis
35	144.5	11.3	169	2 Q9PH31	ARATH Q9ph31 arabidopsis
36	144	11.2	179	2 Q8LPD4	LINUS Q8lpd4 linum usita
37	143.5	11.2	179	2 Q84NG9	VITVI Q84ng9 vitis vinif
38	141.5	11.0	138	2 Q8H2B8	ANAO C Q8h2b8 anacardium
39	141	11.0	126	2 Q8EW54	HELAN Q8ew54 helianthus
40	140.5	11.0	146	1 2SS_BEREX	P04403 bertholleti
41	139	10.9	155	1 2SS2_CAPMA	P30233 capparid ma
42	139	10.9	184	2 Q42444	BRACM Q42444 brassica ca
43	139	10.9	258	1 2SS_RICCO	P01089 ricinus com
44	132.5	10.3	165	2 Q40850	PICGL Q40850 picea glauc
45	132.5	10.3	167	2 Q81411	PICGL Q81411 picea glauc

ALIGNMENTS

RESULT 1

Q647H0_ARAHY
ID Q647H0_ARAHY PRELIMINARY; PRT; 179 AA.
AC Q647H0;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE 2S protein 1 (Fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Yan Y.S., Lin X.D., Zhang Y.S., Wang L., Wu K., Huang S.Z.;
RT "Isolation of peanut genes encoding arachins and conglutins by
RT expressed sequence tags.";
RL Plant Sci. 169:439-445(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Yan Y., Wang L., Huang S.;
RT "cDNA clone of peanut seed storage protein gene.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY722689; AAU21494.1; -, mRNA.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_amyl; 1.
DR SMART; SM00499; AAI; 1.
FT NON_TER 1
SQ SEQUENCE 179 AA; 20850 MW; 68CE6453B489E1DC CRC64;

Alignment Scores:

Pred. No.: 2.99e-65 Length: 179
Score: 842.00 Matches: 157
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 65.7% Indels: 0
DB: 2 Gaps: 0

```

US-10-728-323-2 (1-717) x Q647H0_ARAHY (1-179)
QY 2 CTCACCATAGTACGCTCGCCCTTTCCTCGCTGCGCCACCGCATCTGGAGGCGAG 61
Db 23 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHisAlaSerAlaArgGln 42
QY 62 CAGTGGGAAGCTCAAGGAGACAGAAAGATGCCAGAGCCAGCTCCAGAGGGCGAACCTGAGG 121
Db 43 GlnTrpGluLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 62
QY 122 CCCTGCGAGCAACATCTCATGAGAAAGATCCAAAGCTGACGAGGATTCATATGAACGGGAC 181
Db 63 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 82
QY 182 CCCTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCTATATGATCGGAGCGCTGGA 241
Db 83 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 102
QY 242 TCCTCTCAGCACCAGAGAGGCTGTCATGAGCTGAACGAGTTTGAGACACCAAGG 301
Db 103 SerSerGlnHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 122
QY 302 TGCATGTGCGAGGATTCGACAGATCATGAGAACCCAGAGCGATAGGTTGCAAGGGAGG 361
Db 123 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 142
QY 362 CAACAGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAGTCGCGCTTAGG 421
Db 143 GlnGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 162
QY 422 GCACCCAGCGTTGCGACTTGCGAGCTGCGAAGTGGCGGCAGACAGATAC 472
Db 163 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 179

RESULT 2
Q6PSU1_ARAHY
ID Q6PSU1_ARAHY PRELIMINARY; PRT; 175 AA.
AC Q6PSU1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Seed storage protein SSP2 (Fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Yan Y., Wang L., Liao B., Li H., Lin X., Huang S.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY581854; AAT00599.1; -; mRNA.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_aml; 1.
DR SMART; SM00499; AAI; 1.
FT NON TER
SQ SEQUENCE 175 AA; 20463 MW; B3FB5806B70733DE CRC64;

Alignment Scores:
Pred. No.: 6.7e-65 Length: 175
Score: 838.00 Matches: 156
Percent Similarity: 99.4% Conservative: 0
Best Local Similarity: 99.4% Mismatches: 1
Query Match: 65.4% Indels: 0
DB: 2 Gaps: 0

US-10-728-323-2 (1-717) x Q6PSU1_ARAHY (1-175)
QY 2 CTCACCATAGTACGCTCGCCCTTTCCTCGCTGCGCCACCGCATCTGGAGGCGAG 61
Db 19 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHisAlaSerAlaArgGln 38

```

```

QY 62 CAGTGGGAAGCTCAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAGG 121
Db 39 GlnTrpGluPheGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 58
QY 122 CCCTGCGAGCAACATCTCATGAGAAAGATCCAAAGCTGACGAGGATTCATATGAACGGGAC 181
Db 59 ProCysGluGlnHisLeuMetGlnLysIleGlnArgAspGluAspSerTyrGluArgAsp 78
QY 182 CCCTACAGCCCTAGTCAGGATCCGTACAGCCCTAGTCCTATATGATCGGAGAGCGCTGGA 241
Db 79 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgGlyAlaGly 98
QY 242 TCCTCTCAGCACCAGAGAGGCTGTCATGAGCTGAACGAGTTTGAGACACCAAGG 301
Db 99 SerSerGlnHisGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 118
QY 302 TGCATGTGCGAGGATTCGACAGATCATGAGAACCCAGAGCGATAGGTTGCAAGGGAGG 361
Db 119 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnGlyArg 138
QY 362 CAACAGAGCAACAGTTCAAGAGGAGCTCAGGAACCTTGCTCAACAGTCGCGCTTAGG 421
Db 139 GlnGlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 158
QY 422 GCACCCAGCGTTGCGACTTGCGAGCTGCGAAGTGGCGGCAGACAGATAC 472
Db 159 AlaProGlnArgCysAspLeuAspValGluSerGlyArgAspArgTyr 175

RESULT 3
Q941R0_ARAHY
ID Q941R0_ARAHY PRELIMINARY; PRT; 156 AA.
AC Q941R0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Allergen II (Fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE-21192158; PubMed-11295663; DOI=10.1067/mai.2001.113522;
Viquez O.M., Summer C.G., Dodo H.W.;
RT "Isolation and molecular characterization of the first genomic clone
of a major peanut allergen, Ara h 2."
J. Allergy Clin. Immunol. 107:713-717(2001).
DR EMBL; AY007229; AAK96887.1; -; Genomic_DNA.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_aml; 1.
DR SMART; SM00499; AAI; 1.
FT NON TER
SQ SEQUENCE 156 AA; 18167 MW; 664D6832F91F36D1 CRC64;

Alignment Scores:
Pred. No.: 3.1e-63 Length: 156
Score: 819.00 Matches: 153
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 63.9% Indels: 0
DB: 2 Gaps: 0

US-10-728-323-2 (1-717) x Q941R0_ARAHY (1-156)
QY 2 CTCACCATAGTACGCTCGCCCTTTCCTCGCTGCGCCACCGCATCTGGAGGCGAG 61
Db 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaHisAlaSerAlaArgGln 23
QY 62 CAGTGGGAAGCTCAAGGAGACAGAAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTGAGG 121
Db 24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43

```

QY 122 CCTCGAGCAACATCTCATGAGAGATCCACGTCGAGGATTCATATCAACGGGAC 181
 |||||
 Db 44 ProCysGluGlnHisLeuMetGlnYsIleGlnArgAspGluAspSerTyrGluArgAsp 63
 |||||
 QY 182 CCGTACAGCCCTAGTCAGATCCGTCACGCTAGTCCATATGATCGGAGGCGCTGGA 241
 |||||
 Db 64 ProTyrSerProSerGlnAspProTyrSerProSerProTyrAspArgArgGlyAlaGly 83
 |||||
 QY 242 TCCTCTCAGCACCACAGAGGCTTCGAATGAGCTGAACGAGTTTCAGAACACCAAGG 301
 |||||
 Db 84 SerSerGlnHisGlnArgCysCysAsnGlnLeuAsnGlnPheGluAsnGlnArg 103
 |||||
 QY 302 TGCATGTCGAGGCAATTCACACAGATCATCGAGAACACGAGCGATAGGTTGCGAGGGAGG 361
 |||||
 Db 104 CysMetCysGluAlaLeuGlnGlnIleMetGluAsnGlnSerAspArgLeuGlnArg 123
 |||||
 QY 362 CACAGGAGCAAGTTCAAGAGGAGCTCAGAACTTCGCTCAACAGTGCAGCTTAGG 421
 |||||
 Db 124 GlnGlnGlnGlnPheLysArgGluLeuArgAsnLeuProGlnGlnCysGlyLeuArg 143
 |||||
 QY 422 GCACACAGCGTTGCGACTTGGAGCTGCAAGTGGCGGC 460
 |||||
 Db 144 AlaProGlnArgCysAspLeuAspValGluSerGlyGly 156
 |||||

RESULT 4

Q8GV20_ARAHY
 ID Q8GV20_ARAHY PRELIMINARY; PRT; 172 AA.
 AC Q8GV20;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Allergen Axa h 2.02.
 OS Arachis hypogaea (Peanut).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
 OC Arachis.
 OX NCBI_TaxID=3818;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Chatei J.-M., Bernard H., Orson F.M.;
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY158467; AA07576.1; -; mRNA.
 DR InterPro; IPR003612; AAI.
 DR Pfam; PF00234; Tryp_alpha_aml1; 1.
 DR SMART; SM00499; AAI; 1.
 SQ SEQUENCE 172 AA; 20114 MW; B8BB91C8D8C143AB CRC64;

Alignment Scores:
 Pred. No.: 7.03e-63 Length: 172
 Score: 815.00 Matches: 155
 Percent Similarity: 92.3% Conservatives: 1
 Best Local Similarity: 91.7% Mismatches: 1
 Query Match: 63.6% Indels: 12
 DB: 2 Gaps: 1

US-10-728-323-2 (1-171) x Q8GV20_ARAHY (1-172)

QY 2 CTCACCATCTAGTACCCCTCGCCCTTCTCTCTCGTCCACGCACTCGCGAGGAC 61
 |||||
 Db 4 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 23
 |||||
 QY 62 CAGTGGGAACTCCAGGACAGAGATCCAGAGCCAGCTCGAGGGCGCACTGAGG 121
 |||||
 Db 24 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 43
 |||||
 QY 122 CCTCGGAGCAACATCTCATGAGAGATCCAACTGACGAGGATTCATATCAACGGGAC 181
 |||||
 Db 44 ProCysGluGlnHisLeuMetGlnYsIleGlnArgAspGluAspSerTyrGlyArgAsp 63
 |||||
 QY 182 CCGTACAGCCCTAGTTCAG-----GATCCG 205
 |||||

Db 64 ProTyrSerProSerGlnAspProTyrSerProSerGlnAspProAspArgArgAspPro 83
 |||||
 QY 206 TACAGCCCTAGTCCATATGATCGGAGAGCGCGTGTGATCTCTCAGCACCAACAGAGGTCT 265
 |||||
 Db 84 TyrSerProSerProTyrAspArgArgGlyAlaGlySerSerGlnHisGlnGluArgCys 103
 |||||
 QY 266 TGCATAGCTGAACGAGTTTGAAGAACCAACAAAGTGCATGTGCGAGGCATTGCAACAG 325
 |||||
 Db 104 CysAsnGluLeuAsnGlnPheGluAsnGlnArgCysMetCysGluAlaLeuGlnGln 123
 |||||
 QY 326 ATCATGGAGAACACAGAGCGATAGTTGCAGGGGAGGCAACAGGACCAAGTTCAAGAGG 385
 |||||
 Db 124 IleMetGluAsnGlnSerAspArgLeuGlnGlyArgGlnGlnGlnPheLysArg 143
 |||||
 QY 386 GAGCTCAGGAACCTTCCTCAACAGTGCAGCGCTTAGCGCACACACAGCTTGCGACTTGGAC 445
 |||||
 Db 144 GluLeuArgAsnLeuProGlnGlnCysGlyLeuArgAlaProGlnArgCysAspLeuGlu 163
 |||||

RESULT 5

Q6PSU2_ARAHY
 ID Q6PSU2_ARAHY PRELIMINARY; PRT; 187 AA.
 AC Q6PSU2;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DE Seed storage protein SSP1 (Fragment).
 OS Arachis hypogaea (Peanut).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
 OC Arachis.
 OX NCBI_TaxID=3818;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Yan Y.S., Lin X.D., Zhang Y.S., Wang L., Wu K., Huang S.Z.;
 RT "Isolation of peanut genes encoding arachins and conglutins by
 expressed sequence tags.";
 RL Plant Sci. 169:439-445(2005).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Yan Y., Wang L., Huang S.;
 RT "cDNA clone of peanut seed storage protein gene.";
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY581853; AAT00598.1; -; mRNA.
 DR InterPro; IPR003612; AAI.
 DR Pfam; PF00234; Tryp_alpha_aml1; 1.
 DR SMART; SM00499; AAI; 1.
 FT NON_TER 1
 SQ SEQUENCE 187 AA; 21786 MW; 850AF22C94983F88 CRC64;

Alignment Scores:
 Pred. No.: 7.08e-63 Length: 187
 Score: 815.00 Matches: 155
 Percent Similarity: 92.3% Conservatives: 1
 Best Local Similarity: 91.7% Mismatches: 1
 Query Match: 63.6% Indels: 12
 DB: 2 Gaps: 1

US-10-728-323-2 (1-171) x Q6PSU2_ARAHY (1-187)

QY 2 CTCACCATCTAGTACCCCTCGCCCTTCTCTCTCGTCCACGCACTCGCGAGGAC 61
 |||||
 Db 19 LeuThrIleLeuValAlaLeuAlaLeuPheLeuLeuAlaAlaHisAlaSerAlaArgGln 38
 |||||
 QY 62 CAGTGGGAACTCCAGGACAGAGATCCAGAGCCAGCTCGAGGGCGCACTGAGG 121
 |||||
 Db 39 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 58
 |||||
 QY 122 CCTCGGAGCAACATCTCATGAGAGATCCAACTGACGAGGATTCATATCAACGGGAC 181
 |||||

```

Db      59 ProCysGluGlnHisLeuMetGlnLysIleGlnArgaspGluaspSerTyrGlyArgasp 78
      |||
QY      182 CCGTACAGCCCTAGTCAG-----GATCCG 205
      |||
Db      79 ProTyrSerProSerGlnAspProTyrSerProSerGlnAspProAspArgaspPro 98
      |||
QY      206 TACAGCCCTAGTCATATATCGGAGAGCGCTGGATCCTCTCAGCACCAAGAGAGGTGT 265
      |||
Db      99 TyrSerProSerProTyrAspArgGlyAlaGlySerSerGlnHisGlnGluArgCys 118
      |||
QY      266 TGCATAGAGCTGAACAGATTGAGAACCAACCAAGGTGCATGCGCAGGCATTGCAACAG 325
      |||
Db      119 CysAsnGluLeuAsnGluPheGluAsnAsnGlnArgCysMetCysGluAlaLeuGlnGln 138
      |||
QY      326 ATCATGGAGAACAGAGCGATAGTTGCAGGGGAGGCAACAGGAGCAACAGTTTCAAGAGG 385
      |||
Db      139 IleMetGluAsnGlnSerAspArgLeuGlnGlyArgGlnGlnGlnGlnPheLysArg 158
      |||
QY      386 GAGCTCAGGAACCTGCTCAACAGTGGCGCTTAGGGCCACCACAGCGTTGCGACTTGGAC 445
      |||
Db      159 GluLeuArgAsnLeuProGlnGlnCysGlyLeuArgAlaProGlnArgCysAspLeuGlu 178
      |||
QY      446 GTCGAAAGTGGCGCAGACAGATAC 472
      |||
Db      179 ValGluSerGlyGlyArgaspArgTyr 187
      |||

```

RESULT 6

```

Q7Y1C0 ARAHY
ID Q7Y1C0 ARAHY PRELIMINARY; PRT; 169 AA.
AC QY1C0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Allelgen Ara h 2 isoform (Fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Becker W.-M., Suhr M., Lindner B., Wicklein D., Lepp U.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY117434; AAM78596.1; -; mRNA.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_amyl; 1.
DR SMART; SM00499; AAI; 1.
FT NON_TER 1
SQ SEQUENCE 169 AA; 19768 MW; 0B14A7ED911F34EF CRC64;

```

```

Alignment Scores:
Pred. No.: 2,89e-62 Length: 169
Score: 808.00 Matches: 154
Percent Similarity: 91.7% Conservatives: 1
Best Local Similarity: 91.1% Mismatches: 2
Query Match: 63.1% Indels: 12
DB: 2 Gaps: 1
US-10-728-323-2 (1-717) x Q7Y1C0_ARAHY (1-169)

```

```

QY      2 CTCACCATAGTAGAGCCCTCGCCCTTTCTCTCTCGTCCCAACGATCTGCGAGGCAG 61
      |||
Db      1 LeuThrIleLeuValAlaProAlaLeuPheLeuAlaAlaHisAlaSerAlaArgGln 20
      |||
QY      62 CAGTGGGAACCTCCAGGAGCAGAGATGCCAGGCCAGCTCCGAGGGCGAACCTCAGG 121
      |||
Db      21 GlnTrpGluLeuGlnGlyAspArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 40
      |||
QY      122 CCCTGCCAGCAACATCTCATGCAGAAATCCAAAGTCCAGCAGGATTCATATGAACGGGAC 181
      |||
Db      41 ProCysGluGlnHisLeuMetGlnLysIleGlnArgaspGluaspSerTyrGlyArgasp 60
      |||

```

```

QY      182 CCGTACAGCCCTAGTCAG-----GATCCG 205
      |||
Db      61 ProTyrSerProSerGlnAspProTyrSerProSerGlnAspProAspArgaspPro 80
      |||
QY      206 TACAGCCCTAGTCATATATCGGAGAGCGCTGGATCCTCTCAGCACCAAGAGAGGTGT 265
      |||
Db      81 TyrSerProSerProTyrAspArgGlyAlaGlySerSerGlnHisGlnGluArgCys 100
      |||
QY      266 TGCATAGAGCTGAACAGATTGAGAACCAACCAAGGTGCATGCGCAGGCATTGCAACAG 325
      |||
Db      101 CysAsnGluLeuAsnGluPheGluAsnAsnGlnArgCysMetCysGluAlaLeuGlnGln 120
      |||
QY      326 ATCATGGAGAACAGAGCGATAGTTGCAGGGGAGGCAACAGGAGCAACAGTTTCAAGAGG 385
      |||
Db      121 IleMetGluAsnGlnSerAspArgLeuGlnGlyArgGlnGlnGlnGlnPheLysArg 140
      |||
QY      386 GAGCTCAGGAACCTGCTCAACAGTGGCGCTTAGGGCCACCACAGCGTTGCGACTTGGAC 445
      |||
Db      141 GluLeuArgAsnLeuProGlnGlnCysGlyLeuArgAlaProGlnArgCysAspLeuGlu 160
      |||
QY      446 GTCGAAAGTGGCGCAGACAGATAC 472
      |||
Db      161 ValGluSerGlyGlyArgaspArgTyr 169
      |||

```

RESULT 7

```

Q84TU1 ARAHY
ID Q84TU1 ARAHY PRELIMINARY; PRT; 166 AA.
AC Q84TU1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Seed storage protein (Fragment).
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Seed;
RX PubMed=12582692;
RA Paik-Ro O.G., Seib J.C., Smith R.L.;
RL "Seed-specific, developmentally regulated genes of peanut.";
RL Theor. Appl. Genet. 104:236-240(2002).
DR EMBL; AF366560; AAC61750.1; -; mRNA.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_amyl; 1.
DR SMART; SM00499; AAI; 1.
FT NON_TER 1
FT NON_TER 166
SQ SEQUENCE 166 AA; 19321 MW; BD04F7F26CE7B437 CRC64;

```

```

Alignment Scores:
Pred. No.: 7,35e-61 Length: 166
Score: 792.00 Matches: 151
Percent Similarity: 92.1% Conservatives: 1
Best Local Similarity: 91.5% Mismatches: 1
Query Match: 61.8% Indels: 12
DB: 2 Gaps: 1
US-10-728-323-2 (1-717) x Q84TU1_ARAHY (1-166)

```

```

QY      2 CTCACCATAGTAGAGCCCTCGCCCTTTCTCTCTCGTCCCAACGATCTGCGAGGCAG 61
      |||
Db      2 LeuThrIleLeuValAlaLeuAlaLeuPheLeuAlaAlaHisAlaSerAlaArgGln 21
      |||
QY      62 CAGTGGGAACCTCCAGGAGCAGAGATGCCAGGCCAGCTCCGAGGGCGAACCTCAGG 121
      |||
Db      22 GlnTrpGluLeuGlnGlyAspArgArgCysGlnSerGlnLeuGluArgAlaAsnLeuArg 41
      |||
QY      122 CCCTGCCAGCAACATCTCATGCAGAAATCCAAAGTCCAGCAGGATTCATATGAACGGGAC 181
      |||

```

Db 42 ProCysGluGlnHisLeuMetGlnYsIleGlnArgAspGluAspSerTyrGlyArgAsp 61
 QY 182 CCGTACAGCCCTAGTCAG-----GATCCG 205
 Db 62 ProTyrSerProSerGlnAspProTyrSerProSerGlnAspProAspArgAspPro 81
 QY 206 TACAGCCCTAGTCATATGATCGGAGAGCGGTGGATCCTCTCAGACCAAGAGAGGTGT 265
 Db 82 TyrSerProSerProTyrAspArgArgGlyAlaGlySerSerGlnHisGlnGluArgCys 101
 QY 266 TGCATGAGCTGACAGCTTGTGAAACAACCAAGAGTGTGATGTCGAGGCGATTCGAACAG 325
 Db 102 CysAsnGluLeuAsnGluPheGluAsnAsnGlnArgCysMetCysGluAlaLeuGlnGln 121
 QY 326 ATCATGAGCAACAGAGCGATAGTTGTCAGGGGAGGCAACAGGAGCAACAGTTCAGAGG 385
 Db 122 IleMetGluAsnGlnSerAspArgLeuGlnGlyArgGlnGlnGlnGlnPheYsArg 141
 QY 386 GAGCTCAGGAACCTGCTCAACAGAGTCGGCGCTTAGGGCCACACAGCGTTGCGACTTTGGAC 445
 Db 142 GluLeuArgAsnLeuProGlnGlnCysGlyLeuArgAlaProGlnArgCysAspLeuGlu 161
 QY 446 GTCGAAAGTGGCGGC 460
 Db 162 ValGluSerGlyGly 166

RESULT 8

Q647G9_ARAHY ID Q647G9_ARAHY PRELIMINARY; PRT; 145 AA.
 AC Q647G9;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Conglutinin.
 OS Arachis hypogaea (Peanut).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
 OC Arachis.
 OX NCBI_TaxID=3818;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Yan Y.S., Lin X.D., Zhang Y.S., Wang L., Wu K., Huang S.Z.;
 RT "Isolation of peanut genes encoding arachins and conglutins by
 RT expressed sequence tags";
 RL Plant Sci. 169:439-445(2005).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Yan Y., Wang L., Huang S.;
 RT "cDNA clone of peanut seed storage protein gene.";
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RA Yan Y., Wang L., Huang S.;
 RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY722690; AAU21495.1; -; mRNA.
 DR EMBL; AY849314; AAW32558.1; -; mRNA.
 DR InterPro; IPR003612; AAI.
 DR Pfam; PF00234; Tryp_alpha_aml1; 1.
 DR SMART; SM00499; AAI; 1.
 SQ SEQUENCE 145 AA; 16920 MW; FA3D6FC41CB6267D CRC64;

Alignment Scores:
 Pred. No.: 3.31e-31 Length: 145
 Score: 454.50 Matches: 93
 Percent Similarity: 69.7% Conservative: 15
 Best Local Similarity: 60.0% Mismatches: 30
 Query Match: 35.5% Indels: 17
 DB: 2 Gaps: 4

US-10-728-323-2 (1-717) x Q647G9_ARAHY (1-145)

QY 5 ACCATACTAGTAGCCCTCGCCCTTTCTCTCGTCGCCACGCACTCTGCG---AGGCAG 61
 Db ThrilleLeuValAlaLeuLeuAlaLeuValAlaHisAlaSerAlaMetArgArg 24
 QY 62 CAGTGGGAACCTCCAAGGAGAC---AGAAGATGCCAGAGCCAGCTCGAGAGGGCGAACCTG 118
 Db GluArgGlyArgGlnGlyAspSerSerCysGluArgGlnValAspArgValAsnLeu 44
 QY 119 AGGCCTCGAGCAACATCTCATGCAGAAAGATCCAACGTGACGAGGATTCATATGAACGG 178
 Db LysProCysGluGlnHisIleMetGlnArgIleMetGlyGluGlnGlnGlnTyrAsp--- 63
 QY 179 GACCCGTACAGCCCTAGTCAGATCCGTACAGCCCTAGTCATATGATCGGAGAGCGCT 238
 Db 64 -----SerTyrAspIleArgSerThr 70
 QY 239 GGATCTCTCAGCACCAAGAGAGGTGTCATGAGCTGAACGAGTTTGAGACAACCAA 298
 Db ArgSerSerAspGlnGlnArgCysCysAspGluLeuAsnGluMetGluAsnThrGln 90
 QY 299 AGGTGCATGTGCGAGGCATTGCAACAGATCATGGAGAACCCAGAGCGATAGGTTGCAGGGG 358
 Db ArgCysMetCysGluAlaLeuGlnIleMetGluAsnGlnCysAspArgLeuGlnAsp 110
 QY 359 AGGCAACAGGAGCAACAGTTCAAGAGGAGCTCAGAACTTGCTCAACAGTCGCGCCTT 418
 Db ArgGlnMetValGlnGlnPheLysArgGluLeuMetAsnLeuProGlnGlnCysAsnPhe 130
 QY 419 AGGCACACAGCGTTGCGACTTGAGCTCGMAAGTGGCGGCGAGA 463
 Db 131 ArgAlaProGlnArgCysAspLeuAspVal---SerGlyGlyArg 144

RESULT 9

Q8W251_ARAHY ID Q8W251_ARAHY PRELIMINARY; PRT; 144 AA.
 AC Q8W251;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Conglutinin (Fragment).
 OS Arachis hypogaea (Peanut).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
 OC Arachis.
 OX NCBI_TaxID=3818;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed=12582692;
 RT "Seed-specific, developmentally regulated genes of peanut.";
 RL Theor. Appl. Genet. 104:236-240(2002).
 DR EMBL; AF366561; AAL37561.1; -; mRNA.
 DR InterPro; IPR003612; AAI.
 DR Pfam; PF00234; Tryp_alpha_aml1; 1.
 DR SMART; SM00499; AAI; 1.
 FT NON_TER 1
 SQ SEQUENCE 144 AA; 16750 MW; B776486138A852C3 CRC64;

Alignment Scores:
 Pred. No.: 3.07e-30 Length: 144
 Score: 443.50 Matches: 90
 Percent Similarity: 69.7% Conservative: 18
 Best Local Similarity: 58.1% Mismatches: 30
 Query Match: 34.6% Indels: 17
 DB: 2 Gaps: 4

US-10-728-323-2 (1-717) x Q8W251_ARAHY (1-144)

QY 5 ACCATACTAGTAGCCCTCGCCCTTTCTCTCGTCGCCACGCACTCTGCG---AGGCAG 61
 Db ThrilleLeuValAlaLeuLeuAlaLeuValAlaHisAlaSerAlaMetArgArg 23


```
Db 42 Asp-:::|||||SerTyrAspIleArg 47
233 GGGCGTGGATCCTCTCAGCACCAAGAGAGGTGTTGCAATGAGCTGAACGAGTTTGAGAAC 292
Db 48 SerThrArgSerSerAspGlnGlnGlnArgCysCysAspGluLeuAsnGluMetGluAsn 67
293 AACCAAGGTGCATGTCGAGGCGATTGCAACAGATCATGGAGAACACAGAGCGATAGTTG 352
Db 68 ThrGlnGlyCysMetCysGluAlaLeuGlnGlnMetGluAsnGlnCysAspArgLeu 87
353 CAGGGAGGCAACAGAGCAACAGTTCAAGAGGAGCTCAGAACTTGCTCAACAGTGC 412
Db 88 GlnAspArgGlnMetValGlnGlnPheLysArgGluLeuMetSerLeuProGlnGlnCys 107
413 GGGCTTAGGCACCAACAGCGTTGCGACTTGGAGTCGAAAGTGGCGGCAGA 463
Db 108 AsnPheArgAlaProGlnArgCysAspLeuAspVal---SerGlyGlyArg 123

RESULT 12
Q9SQG5_ARAHY
ID Q9SQG5_ARAHY PRELIMINARY; PRT; 129 AA.
AC Q9SQG5_ARAHY
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Allergen Arah6 (Fragment).
GN Name=Ara h 6;
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]_TaxID=3818;
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Seed;
RX MEDLINE=99406463; PubMed=10474031; DOI=10.1159/000024203;
RA Kleber-Janke T., Cramer R., Appenzeller U., Schlaak M., Becker W.-M.;
RT "Selective cloning of peanut allergens, including profilin and 2S
RT albumins, by phage display technology.";
RL Int. Arch. Allergy Immunol. 119:265-274(1999).
DR EMBL; AF092846; AAD56337.1; -; mRNA.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; Tryp_alpha_amyl; 1.
DR SMART; SM00499; AAI; 1.
FT NON_TER
SQ SEQUENCE 129 AA; 14929 MW; 4048D3418BB9BE40 CRC64;

Alignment Scores:
Pred. No.: 2,07e-25 Length: 129
Score: 388.50 Matches: 80
Percent Similarity: 66.4% Conservative: 15
Best Local Similarity: 55.9% Mismatches: 31
Query Match: 30.3% Indels: 17
DB: 2 Gaps: 4

US-10-728-323-2 (1-717) x Q9SQG5_ARAHY (1-129)
QY 41 GCCCAGCATCTGCG---AGGCAGCATGGGAACCTCAAGAGAC---AGAAGATGCCAG 94
Db 1 AlaHieAlaSerAlaMetArgArgGluArgGlyArgGlnGlyAspSerSerCysGlu 20
QY 95 AGCCAGCTCGAGAGGCGGACCTGAGGCCCTCGGACACATCTCATGCAGAAGATCCAA 154
Db 21 ArgGlnValAspGlyValAsnLeuLysProCysGluGlnHieSileMetGlnArgIleMet 40
QY 155 CGTGAGGAGGATTTCATATGAACGGGACCCGTACAGCCCTAGTCAGGATCCGTACAGCCCT 214
Db 41 GlyGluGlnGluGlnTyrAsp----- 47
QY 215 AGTCCATATGATCGGAGGCGCTGGATCCTCTCAGCAACCAAGAGAGGTGTTGCAATGAG 274
```

```
Db 48 ---SerTyrAsnPheGlySerThrArgSerSerAspGlnGlnGlnArgCysCysAspGlu 66
275 CTGAACGAGTTTGGAGAACCAACAAAGGTGTCATGTGCGAGGCAATTGCAACAGATCATGGAG 334
Db 67 LeuAsnGluMetGluAsnThrGlnArgCysMetCysGluAlaLeuGlnGlnIleMetGlu 86
335 AACCAAGCGCATAGTTGCGAGGAGGCAACAGGAGCAACAGTTCGAAGGAGGAGTCCAGG 394
Db 87 AsnGlnCysAspGlyLeuGlnAspArgGlnMetValGlnHiePheLysArgGluLeuMet 106
395 AACTTTCCTCAACAGCTGCGGCTTAGGCACCAACAGGAGCAACAGTTCGAAGGAGGAGTCCAGG 454
Db 107 AsnLeuProGlnGlnCysAsnPheGlyAlaProGlnArgCysAspLeuAspVal---Ser 125
455 GCGCGCAGA 463
Db 126 GlyGlyArg 128

RESULT 13
Q9SQH1_ARAHY
ID Q9SQH1_ARAHY PRELIMINARY; PRT; 160 AA.
AC Q9SQH1_ARAHY
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Allergen.
GN Name=Ara h 7;
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]_TaxID=3818;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Virginia; TISSUE=Seed;
RX MEDLINE=99406463; PubMed=10474031; DOI=10.1159/000024203;
RA Kleber-Janke T., Cramer R., Appenzeller U., Schlaak M., Becker W.-M.;
RT "Selective cloning of peanut allergens, including profilin and 2S
RT albumins, by phage display technology.";
RL Int. Arch. Allergy Immunol. 119:265-274(1999).
DR EMBL; AF091737; AAD56719.1; -; mRNA.
DR InterPro; IPR003612; AAI.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 160 AA; 18417 MW; 9F9E4CEE6808D4C CRC64;

Alignment Scores:
Pred. No.: 1.26e-17 Length: 160
Score: 300.00 Matches: 72
Percent Similarity: 61.8% Conservative: 30
Best Local Similarity: 43.6% Mismatches: 33
Query Match: 23.4% Indels: 31
DB: 2 Gaps: 7

US-10-728-323-2 (1-717) x Q9SQH1_ARAHY (1-160)
QY 2 CTCACCATAGTAGCCCTC---GCCCTTTTCTCTCTCGCTGCC----- 43
Db 5 LeuSerIleLeuValAlaLeuLeuGlyAlaLeuLeuValAlaSerAlaThrArgTyr 24
QY 44 -----CACGCATCTCGCAGGAGGAGCGAGTGGGAACCTC-----CAAGGAGACAGAGA 88
Db 25 AspProAspArgGlySerArgGlySerArgTyrAspAlaProSerArgGlyAspGln 44
QY 89 TCCAGAGCCAGCTCGAGAGGCGAACCTGAGGCCCTCGGACCAACATCTCATGCAGAG 148
Db 45 CysGlnArgGlnLeuGlnArgAlaAsnLeuArgProCysGluGluHieMetArgArg 64
QY 149 ATCCAACTGACGAGGATTTCATGAACGGACCGGTACAGCCCTAGTCAGGATCCGTAC 208
Db 65 ValGluGlnGluGlnGluGln---GluGlnAspGluTyr----- 76
QY 209 AGCCCTAGTCCATATGATCGGAGGCGCTGGATCCTCTCTCAG----- 250
```


Qy	265	TACAGCCCTAGTCTCATATGATCGGAGAGCGCTGGATCCTCTACGACCAAGAGAGGTGT	265
Db	80	TyrSerGlu-----GluSerGluGluLeuAspGlnCys	90
Qy	266	TGCAATGAGCTGAACGAGTTTGAGAAACAACAAAGGTGTCATGTGCGAGGCATTGCAACAG	325
Db	91	CysGluGluLeuAsnGluLeu---AsnSerGlnArgCysGlnCysArgAlaLeuGlnGlnGln	109
Qy	326	ATCATGGAGAACCGACGAGCGATGTTGCGAGGGGAGGCAACAGGAGCAACAGTTCTCAAGAGG	385
Db	110	IleTyrGluSerGlnSerGluGlnCysGluGlySerGlnGlnGlnGlnGlnGluGln	129
Qy	386	GAGCTCAGGAACCTTGCTCAACAGATGCGCGCCTTAGGGCCACACAGCGCTCGCATTTGGAC	445
Db	130	GluLeuGluGlyLeuProArgThrCysGlyPheGlyProLeuArgArgCysAspValasn	149

Search completed: May 15, 2006, 22:17:08
Job time : 167.351 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 21:23:02 ; Search time 45.6522 Seconds
(without alignments)
4400.310 Million cell updates/sec

Title: US-10-728-323-3

Perfect score: 2779

Sequence: 1 cggcgcaacgagaggagaa.....ctccgagggtgtggtcttaa 1524

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 486326

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xl

-Q=/abs/ABSWEB/spool/US1078323/runat_15052006.172130.22366/app_query.fasta.1
-DB=A_Geneseq -QFW=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US1078323 @CGN 1.1 605 @runat 15052006.172130.22366 -NCPU=6 -ICPU=3
-NO_WMAP -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A_Geneseq.21:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2673	96.2	510	6	ABU52484 Peanut Ar
2	2673	96.2	510	7	ADG27544 Peanut al
3	2667	96.0	510	4	AAU04708 Anaphylac
4	2665	95.9	507	8	ADM12139 Arachis h
5	2663	95.8	510	2	AAV15246 Peanut al
6	2652	95.4	512	2	AAV40912 Ara h 3 a
7	2638	94.9	526	3	AAU04711 Modified
8	2638	94.9	526	4	AAU04711 Modified
9	2638	94.9	526	4	AAU05036 Modified

10	2615	94.1	507	8	ADO38357	Ado38357	Peanut al
11	1502	54.0	481	5	ABG71266	Abg71266	Glycine m
12	1502	54.0	481	7	ADH89253	Adh89253	G. max gl
13	1502	54.0	481	7	ADL90187	Adl90187	Soybean g
14	1502	54.0	481	8	ADG43988	Adg43988	G. max gl
15	1502	54.0	488	8	ADG74499	Adg74499	Plant ful
16	1502	54.0	491	8	ADX74681	Adx74681	Plant ful
17	1502	54.0	492	8	ADX74470	Adx74470	Plant ful
18	1502	54.0	492	8	ADX76928	Adx76928	Plant ful
19	1499.5	54.0	466	8	ADY24175	Ady24175	Plant ful
20	1499.5	54.0	466	8	ADY24189	Ady24189	Plant ful
21	1499.5	54.0	485	5	ABG71265	Abg71265	Glycine m
22	1499.5	54.0	485	7	ADH89247	Adh89247	G. max gl
23	1499.5	54.0	485	7	ADL90186	Adl90186	Soybean g
24	1499.5	54.0	485	8	ADG43982	Adg43982	G. max gl
25	1499.5	54.0	488	8	ADX76279	Adx76279	Plant ful
26	1499.5	54.0	489	8	ADY24211	Ady24211	Plant ful
27	1499.5	54.0	489	8	ADY24504	Ady24504	Plant ful
28	1499.5	54.0	489	8	ADX76246	Adx76246	Plant ful
29	1499.5	54.0	489	8	ADY24125	Ady24125	Plant ful
30	1499.5	54.0	489	8	ADX75310	Adx75310	Plant ful
31	1499.5	54.0	489	8	ADY24174	Ady24174	Plant ful
32	1499.5	54.0	489	8	ADX87453	Adx87453	Plant ful
33	1499.5	54.0	489	8	ADY24183	Ady24183	Plant ful
34	1499.5	54.0	489	8	ADX88748	Adx88748	Plant ful
35	1499.5	54.0	489	8	ADX87543	Adx87543	Plant ful
36	1499.5	54.0	489	8	ADY24118	Ady24118	Plant ful
37	1499.5	54.0	489	8	ADY24503	Ady24503	Plant ful
38	1499.5	54.0	489	8	ADX74617	Adx74617	Plant ful
39	1499.5	54.0	489	8	ADY24173	Ady24173	Plant ful
40	1499.5	54.0	489	8	ADX95765	Adx95765	Plant ful
41	1498	53.9	489	8	ADX76273	Adx76273	Plant ful
42	1496	53.8	476	9	ADZ47094	Adz47094	Soybean g
43	1496	53.8	477	9	ADZ47108	Adz47108	Soybean m
44	1496	53.8	485	9	ADZ47102	Adz47102	N-termina
45	1496	53.8	485	9	ADZ47106	Adz47106	C-termina

ALIGNMENTS

RESULT 1

ABU52484

ID ABU52484 standard; protein; 510 AA.

XX AC ABU52484;

XX AC ABU52484;

DT 10-MAR-2003 (first entry)

XX DE Peanut Ara h3 protein sequence.

XX DE Peanut Ara h3 protein sequence.

KW Peanut; allergy; Ara h1; Ara h2; Ara h3; IgE binding site;

KW anaphylactic food allergen; antiallergenic; vaccine; wound healing.

XX OS Arachis hypogaea.

XX OS Arachis hypogaea.

XX PN WO200274250-A2.

XX PN WO200274250-A2.

PD 26-SEP-2002.

XX PD 26-SEP-2002.

PF 18-MAR-2002; 2002WO-US009108.

XX PF 18-MAR-2002; 2002WO-US009108.

PR 16-MAR-2001; 2001US-0276822P.

XX PR 16-MAR-2001; 2001US-0276822P.

PR 18-MAR-2002; 2002US-00276822.

XX PR 18-MAR-2002; 2002US-00276822.

XX (PANA-) PANACEA PHARM.

XX (PANA-) PANACEA PHARM.

PI Caplan M, Sosin H, Sampson H, Bannon GA, Burks WA, Cockrell G;

PI Compadre CW, Connaughton C, Helm RM, King NE, Kopper RA, Maleki SJ;

PI Rabinjohn PA, Shin DS, Stanley JS;

XX WPI; 2003-018765/01.

DR WPI; 2003-018765/01.

XX N-PSDB; ABX70612.

XX N-PSDB; ABX70612.

PT New modified anaphylactic food allergen, useful for preventing or
 PT treating allergic reactions associated with e.g. anaphylactic allergens.
 XX
 PS
 XX Claim 27; Fig 68B; 300pp; English.
 CC The invention relates to a modified anaphylactic food allergen has an
 CC amino acid sequence that is substantially identical to that of natural
 CC anaphylactic food allergen, except for a cysteine residue that has been
 CC modified so that it cannot participate in the disulphide bond. The
 CC modification may also comprise mutation of the IgE binding sites to
 CC reduce allergenicity. Also included are: (1) a method of making a
 CC modified anaphylactic food allergen; (2) a nucleotide molecule encoding
 CC or for causing a site specific mutation in the modified anaphylactic food
 CC allergen; (3) a transgenic plant or animal expressing the modified
 CC anaphylactic food allergen; (4) a method of treating an individual by
 CC reducing the clinical response to a natural anaphylactic food allergen;
 CC and an isolated fragment of peanut allergen Ara h 1. The modified
 CC anaphylactic food allergen is useful for preventing or treating allergic
 CC reactions associated with any natural allergen such as food, insect,
 CC rubber or preferably anaphylactic allergens. It is also useful for
 CC treating wounds in mammals such as bovine, canine, feline, caprine,
 CC ovine, porcine, murine or equine species. The present sequence is a
 CC peanut allergen (e.g. Ara h1, h2 or h3)
 XX
 SQ Sequence 510 AA;

Alignment Scores:

Pred. No.: 4,28e-236 Length: 510
 Score: 2673.00 Matches: 507
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 96.2% Indels: 0
 DB: 6 Gaps: 0

US-10-728-323-3 (1-1524) x ABU52484 (1-510)

QY	1	CGCAGCAACCGAGGAGACGGTCCAGTTCAGGGCCCTCAATGGCGAGACCTGAC	60
DB	4	ArgGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	23
QY	61	AATCGCATTCGAATCAGAGGCGGTACATTGAGACTTGGAAACCCCAACACACAGGAGTTC	120
DB	24	AsnArgIleGluSerGluGlyGlyTrpIleGluThrTrpAsnProAsnGlnGluPhe	43
QY	121	GAATGCGCCGGCTCGCCCTCTCTCGCTTAGTCTCGCCGCAACGCCCTTCGTAGCCCT	180
DB	44	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro	63
QY	181	TTCTACTCCCAATGCTCCCGAGGATCTTCATCCAGCAAGGAGGGGATACTTTGGTTG	240
DB	64	PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu	83
QY	241	ATATTCCCTGGTTGTCCTAGACACTATGAAGAGCCTCACACAAAGTTCGTGATCTCAG	300
DB	84	IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgSerGln	103
QY	301	TCCCAAGACACCAAGACCTCTCCAGGAGAGACCAACCAAGCCCAACGACGATAGT	360
DB	104	SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer	123
QY	361	CACGAGAGTGCACCGTTTCGATGAGGTGATCTCATCCAGTTCCTCCCGGTGTGCT	420
DB	124	HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	143
QY	421	TTCTGGCTCTACACGACCAACGACACTGATGTGTGCTGCTGTTCTCTTACTGACCAAC	480
DB	144	PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn	163
QY	481	AACAAACGACCAACGAGTTCAGTTCCTCCAGGAGATTCATATTGGCTGGGAACACGGAG	540
DB	164	AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu	183
QY	541	CAAGAGTTCCTTAAGGTACCACAAACAGACAAAGACGACGAAAGAGCTTACCATAT	600

DB	184	GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgArgSerLeuProTyr	203
QY	601	AGCCCATACAGCCCGCAAGTTCAGCTCAGCAAGAGACGCTGAATTTAGCCCTCCAGGA	660
DB	204	SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly	223
QY	661	CAGCAGCGCGAGAGAACGAGCAGGACAAAGAAAGAAAGAAAGAGGTGGAACATCTTC	720
DB	224	GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe	243
QY	721	ACGGCTTCACCGCCGAGTTCCTGGAACAAGCTTCACAGTTCACGACAGACAGATAGTG	780
DB	244	SerGlyPheThrProGlnPheLeuGlnGlnAlaPheGlnValAspArgGlnIleVal	263
QY	781	CAAAACCTAAGAGCGGAGACCCGAGAGTGAAGAGGGAGCCATTGTGACGTGAGGGGA	840
DB	264	GlnAsnLeuArgGlyGluThrGluSerGluGluGlyAlaIleValThrValArgGly	283
QY	841	GGCTCAGAAATCTTGAGCCCATAGAAAGACACGTGCGCAGCAAGAAAGAGGATACGAT	900
DB	284	GlyLeuArgIleLeuSerProAspArgLysArgAlaAspGluGluGluTyrAsp	303
QY	901	GAAGATGAATATCAATACCATGAAGAGGATAGAAGCGGTGGCAGGGGAAACGAGCAGG	960
DB	304	GluAspGluTyrGluTyrAspGluGluAspArgArgArgGlySerArgGlyArg	323
QY	961	GGCAATGGTATTGAAGAGACGATCTGCACCCGAAGTGTAAAGAAACATTTGTAAGAAC	1020
DB	324	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaAlaLysAsnIleGlyArgAsn	343
QY	1021	AGATCCCTTGACATCTCAACCTCAGCTGTTCACTCAAAACTGCCAACGATCTCAAC	1080
DB	344	ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn	363
QY	1081	CTTCTAATTAATAGTGGCTTGGACCTAGTGTGTAATATGGAATCTCTACAGGAATGCA	1140
DB	364	LeuLeuLeuLeuArgTyrLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla	383
QY	1141	TTGTTTGTCTGCTCACTACAACACCAACCGCACACAGCATCATATATCGATTGAGGGACGG	1200
DB	384	LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg	403
QY	1201	GCTCAGCTCGAAGTCTGTGACAGCAACCGCAACAGAGTGTACGACGAGGAGTTCAGAG	1260
DB	404	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluLeuGlnGlu	423
QY	1261	GGTCACGTGCTTGTGTGTCACAGAACTTCGCGCTCGCTGGAAAGTCCGACGACGAGAAC	1320
DB	424	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn	443
QY	1321	TTCAATACGTGGCATTCAGACAGACTCAAGCCCGCAGCATAGCCCACTCGCCGCTGAA	1380
DB	444	PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	463
QY	1381	AACTCCGTCTAGATAAACCCTCGCGAGGAGTGGTTGCAAAATTCATATGSCCTCCAAAGG	1440
DB	464	AsnSerValIleAspAsnLeuProGluGluValAlaAsnSerTyrGlyLeuGlnArg	483
QY	1441	GAGCAGCGAAGGCGCTTAAGAAACAACCCCTTCAAGTTCTTTCGTTCCACCGCTCTCAG	1500
DB	484	GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProProSerGln	503
QY	1501	CAGTCTCCGAGGCTGTGGCT 1521	
DB	504	GlnSerProArgAlaValAla 510	

RESULT 2

ADG27544
 ID ADG27544 standard; protein; 510 AA.

XX

AC ADG27544;

XX

181 TTCTACTCCAAATGCTCCCAAGAGATCTTCATCCAGCAAGGAGGATCTTTGGTTG 240
Db PheTyr-SerAsnAlaProGlnGlnIlePheIleGlnGlyArgGlyTyrPheGlyLeu 83
241 ATATTCCTGTTGCTCTAGACACTATGAGAGCCTCACACAAAGTCTGTCATCTCAG 300
Db IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgSerGln 103
301 TCCCAAGACCAACCAAGACGCTCTCCAAGGAGAGAGACCAAGCCAAACAGCAACGAGATAGT 360
Db SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 123
361 CACAGAAGGTGCACCGTTTCGATGAGGGTGTCTCATTCAGTCCCAACCGGTGTGCT 420
Db HisGlnIysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143
421 TTCTGGCTCTACAAACACGACACTGTATGTTGTTGCTGTTCTTCTACTGACACCAAC 480
Db PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 163
481 AACCAACGACAAACAGCTTGATCAGTTCCCGAGAGATTCATATTTGGTGGGAACACGGAG 540
Db AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 183
541 CAAGAGTTCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAGACTTACCATAT 600
Db GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 203
601 AGCCCATACAGCCCGCANAGTCAGCTTAGACAGAAGCGGTGAATTTAGCCCTCCAGGA 660
Db SerProTyr-SerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 223
661 CAGCACAGCCGCGAGAGAACAGCAGGACAAAGAAAGAAACCAAGGTGGAAACATCTTC 720
Db GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe 243
721 AGCGGCTTACGCGGAGTTCTCGAACCAAGCTTCCAGGTTCCAGACAGACAGATAGTG 780
Db SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 263
781 CAAACCTTAAGCGCAGACCGAGAGTGAAGAGAGGGAGCCATTTGTGACGTGAGGGA 840
Db GlnAsnLeuArgGlyGluThrGlnSerGluGluGluGlyAlaIleValThrValArgGly 283
841 GGCCTCAGATCTTGAGCCACAGATAGAAAGACGTGCCGACCAAGAGAGGAATACGAT 900
Db GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrAsp 303
901 GAAGATGAATATGAATACGATGAAGAGATAGAGCGGTGGCAGGGAAACGAGCGCAGG 960
Db GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyThr 323
961 GCGAATGGTATTGAAGAGAGATCTGCCCGCAAGTGTAAAGAGACATTGTGTAGAAC 1020
Db GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysAsnIleGlyArgAsn 343
1021 AGATCCCTCAGATCTACAAACCTCAAGCTGGTTCACTCAAACTGCCAACGATCTCAAC 1080
Db ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 363
1081 CTCTTAATCTAGTGGCTGGACCTAGTGTGAATATGGAAATCTCTACAGGAATGCA 1140
Db LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 383
1141 TTCTTTGCTCGCTCACTACACCAACGACACAGCATCATATATCATGTAGGGGACCG 1200
Db LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 403
1201 GCTCAGCTGCAAGTCTGAGCAGCAACCGCAACAGAGTGTACCACGAGAGCTTCAAGAG 1260
Db AlaHisValGlnValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 423
1261 GGTACGCTGTTGTGGTCCACAGAACTTCGCGCTCGCTGGAAAGTCCCGAGCGAGAAC 1320

424 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 443
1321 TTTCGAATACGTGGCATTCGAAGACAGACTCAAGCCCGCAGCATAGCCAACCTCGCCGCTGAA 1380
Db PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 463
1381 AACTCCGTCATAGATAACCTGCCGAGGAGGTGGTTGCAAAATTCATATGCGCTCCAAAGG 1440
Db AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 483
1441 GAGCAGGCGAAGCAGCTTAAGAACAAACACCCCTTCAAGTCTTCGTTCCACGCTCTCAG 1500
Db GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProProSerGln 503
1501 CAGTCTCCGAGGCGTGTGGCT 1521
Db GlnSerProArgAlaValAla 510
RESULT 4
ADM12139
ID ADM12139 standard; protein; 507 AA.
XX
AC ADM12139;
XX
DT 20-MAY-2004 (first entry)
XX
DE Arachis hypogaea 3 (Ara h3) protein.
XX
KW antigen presentation enhancing hybrid polypeptide; mammalian Ii-Key;
KW MHC class II; antibacterial; virucide; fungicide; antirheumatic;
KW antiarthritic; neuroprotective; dermatological; immunosuppressive;
KW antiinflammatory; antidiabetic; antithyroid; immune;
KW rheumatoid arthritis; multiple sclerosis; lupus erythematosus;
KW diabetes mellitus; myasthenia gravis; autoimmune thyroiditis;
KW scleroderma; dermatomyositis; pemphigus.
XX
OS Arachis hypogaea.
XX
PN US2003235594-A1.
XX
PD 25-DEC-2003.
XX
PF 17-SEP-2002; 2002US-00245871.
XX
PR 14-SEP-1999; 99US-00396813.
PR 17-JUL-2002; 2002US-00197000.
XX
PA (ANTI-) ANTIGEN EXPRESS INC.
XX
PI Humphreys R, Xu M;
XX
DR WPI; 2004-070554/07.
XX
PT Novel Ii-Key/antigen presentation enhancing hybrid polypeptide, useful
PT for treating infections, rheumatoid arthritis, multiple sclerosis, lupus
PT erythematosus and diabetes mellitus.
XX
PS Example 3; Page 21; 87pp; English.
XX
CC The invention relates to a novel antigen presentation enhancing hybrid
CC polypeptide. The novel polypeptide has an N-terminal element consisting
CC of 4-16 residues of a mammalian Ii-Key peptide and its non-N-terminal
CC deletion modifications, a chemical structure covalently linking the N-
CC terminal element to an MHC class II-presented epitope of a C-terminal
CC element. The C-terminal element comprises an antigenic epitope, which
CC binds to an antigenic peptide binding site of an MHC class II molecule.
CC The antigen presentation enhancing hybrid polypeptide has the following
CC activities: antibacterial, virucide, fungicide, antirheumatic,
CC antiarthritic, neuroprotective, dermatological, immunosuppressive,
CC antiinflammatory, antidiabetic, and antithyroid. The antigen presentation
CC enhancing hybrid polypeptide is useful for modulating the immune response
CC in an individual and for treating infections (such as bacteria, virus,

CC parasite and fungus), rheumatoid arthritis, multiple sclerosis, lupus
CC erythematous, diabetes mellitus, myasthenia gravis, autoimmune
CC thyroiditis, scleroderma, dermatomyositis and pemphigus. This sequence
CC represents a mammalian II key related protein of the invention.
XX
SQ Sequence 507 AA;

Alignment Scores:
Pred. No.: 2,32e-235 Length: 507
Score: 2665.00 Matches: 506
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 95.9% Indels: 0
DB: 8 Gaps: 0

US-10-728-323-3 (1-1524) x ADM12139 (1-507)

QY	1	CGGCAGCAACCGGAGGACGGTGCAGTTCAGCGCCTCAATGCGCAGACCTGAC	60
DB	1	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	20
QY	61	AATCGCATTAATCAGAGGGCGTTACATTGAGACTTGGAAACCCCAACACAGGAGTTC	120
DB	21	AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe	40
QY	121	GAATGCGCGCGTGGCCTCTCTCGTTCAGTTCCTCGCGCGCAACCGCCTTCGTAGCCT	180
DB	41	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro	60
QY	181	TTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGATACTTTGGTTG	240
DB	61	PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu	80
QY	241	ATATTCCTCGTTGTCTCAGACACTATGAAGAGCCTCACACAAAGTCTCGATCTCAG	300
DB	81	IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgSerGln	100
QY	301	TCCCAAGACCCACAGACGCTTCCAAGGAGAGACCAAGCCCAACAGCAACAGATAGT	360
DB	101	SerGlnArgProProArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer	120
QY	361	CACCAGAGGTGCACCGTTTCGATGAGGGTGATCTCATTCAGTCCACCGGTGTGCT	420
DB	121	HisGlnLysValHisArgThrAspGluGlyAspLeuIleAlaValProThrGlyValAla	140
QY	421	TTCTGGCTCTAACCGACCACTGATGTTGTTGCTGTTCTCTCTACTGACCAAC	480
DB	141	PheTrpLeuTyrAsnAspHisAspThrAspValAlaValSerLeuThrAspThrAsn	160
QY	481	AACACGACCAACCGCTTGATCAGTTCGCCAGGAGATTCAATTTGGCTGGGAACCGGAG	540
DB	161	AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu	180
QY	541	CAAGAGTTCCTAAGGTACCAGCAACAAGCAGACAAAGCAGACAAAGAGCTTACCATAT	600
DB	181	GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr	200
QY	601	AGCCCATACAGCCCGCAAGTCAAGCTTAGACAAGAGCGTGAATTTAGCCCTCGAGGA	660
DB	201	SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly	220
QY	661	CAGCAGCGCCAGAGACCGAGCAGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	720
DB	221	GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluGluGluGlyAsnIlePhe	240
QY	721	AGCGGCTTCACGCGGAGTTCCTGGAAACAAGCCTTCCAGGTTGACGACAGATAGTG	780
DB	241	SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal	260
QY	781	CAAAACCTAAGCGCAGACCGAGAGTGAAGAGAGGAGGAGCCATTGTGACAGTGAGGGA	840
DB	261	GlnAsnLeuArgGlyGlyThrGluSerGluGluGluGlyAlaIleValThrValArgGly	280

QY	841	GGCCTCAGAACTCTGAGCCAGATAGAAAGAGACGCTGCCGACGAAGAAGAGGAATACGAT	900
DB	281	GlyLeuArgIleLeuSerProAspArgLysArgAlaAspGluGluGluGluLysArg	300
QY	901	GAAGATGAATATCAATACGATGAAGAGGATAGAAGCGGTGGCAGGGGAACGAGCAGG	960
DB	301	GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg	320
QY	961	GGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGTGTAATAAAGAACATTTGGTGAAC	1020
DB	321	GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn	340
QY	1021	AGATCCCTGACATCTACAAACCTCAAGCTGTGTTCACTCAAACTCCCAACGATCTCAAC	1080
DB	341	ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn	360
QY	1081	CTTCTAATATCTAGTGGCTTGGACCTAGTGTCTGAATATGGAATCTCTACAGGAATGCA	1140
DB	361	LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla	380
QY	1141	TTGTTTGTCTCCTACATAACACCAACGACACAGCATCATATATCGATTGAGGGACCG	1200
DB	381	LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg	400
QY	1201	GCTCAGCTCAAGTCTGGACAGCAACGCAACAGAGTGTACGACGAGGAGCTTCAAGAG	1260
DB	401	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu	420
QY	1261	GCTCAGCTCTTGTGTGTCACAGAACTTTCGCCGTCTGGAAGTCCACAGACGAGAAC	1320
DB	421	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn	440
QY	1321	TTGCAATAGTGGCATTCAGACAGACTCAAGCCCGACAGATAGCCACCTCCGCGGTGAA	1380
DB	441	PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	460
QY	1381	AACTCCGTTCATAGATAACCTGCCGAGGAGGTGGTTGCAAAATTCATATGGCCTCCAAAGG	1440
DB	461	AsnSerValIleAspAsnLeuProGluGluValAlaAlaAsnSerTyrGlyLeuGlnArg	480
QY	1441	GAGCAGCAAGGAGGAGCTTAAGAACCAACACCCCTTCAAGTTCCTTCCACCGTCTCAG	1500
DB	481	GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPheValProProSerGln	500
QY	1501	CAGTCTCCGAGCGCTGGCT 1521	
DB	501	GlnSerProArgAlaValAla 507	
RESULT 5			
ID	AA15246	standard; protein; 510 AA.	
AC	AA15246;		
DT	17-OCT-2003 (revised)		
DT	09-NOV-1999 (first entry)		
XX			
DE	Peanut allergen, Ara h 3, amino acid sequence.		
KW	allergy; immune response; transgenic; allergen; epitope;		
KW	immunoglobulin E; Ig E; binding site; peanut.		
OS	Arachis hypogaea.		
XX			
PN	WO9938978-A1.		
PD			
XX	05-AUG-1999.		
PF	29-JAN-1999; 99WO-US002031.		
XX			
PR	31-JAN-1998; 98US-0073283P.		
PR	13-FEB-1998; 98US-0074590P.		
PR	13-FEB-1998; 98US-0074624P.		


```

XX AAY40912;
AC
XX
XX 17-OCT-2003 (revised)
DT
XX 06-DEC-1999 (first entry)
DT
XX Ara h 3 allergen sequence.
DE
XX Peanut; allergen; Ara H 1; IgE; immunoglobulin E; epitope; Ara h 3;
KW allergic reaction; ss.
KW
XX Arachis hypogaea.
OS
XX
XX Key Location/Qualifiers
FH 187..188
FT Region
FT
FT /note= "these two amino acid residues are not indicated
FT in the Ara h 3 sequence provided in the sequence listing
FT (seq ID No: 6) and Fig 51 of the specification"
FT Misc-difference 374
FT /note= "encoded by CCT"
FT
XX WO9945961-A1.
PN
XX
XX 16-SEP-1999.
PD
XX
XX 12-MAR-1999; 99WO-US005494.
PF
XX
XX 12-MAR-1998; 98US-0077763P.
PR
XX 11-MAR-1999; 99US-00077763.
PR
XX (UYAR-) UNIV ARKANSAS.
PA
XX
XX Burks W, Helm RM, Cockrell G, Bannon GA, Stanley JS, Shin DS;
PI Sampson H, Compadre CM, Huang SK, Maleki SJ, Kopper RA;
PI
XX
XX WPI; 1999-551218/46.
DR
XX N-PSDB; AAZ22280.
DR
XX
XX Tertiary structure of peanut allergen Ara h 1 for protection of a host
PT animal from allergic reaction.
PT
XX Claim 8; Fig 11; 193pp; English.
PS
XX
XX The invention provides a tertiary structure for the peanut allergen Ara H
CC 1. The Ara H 1 allergen is found to contain 23 linear IgE-binding
CC epitopes. The invention also provides an isolated recombinant peanut
CC allergen designated Ara h 3 and a nucleotide molecule encoding the peanut
CC allergen Ara h 3. Molecules of the invention are used to protect a host
CC animal from allergic reaction, particularly using a modified allergen
CC which is less reactive with IgE. The invention may also be used to ensure
CC that the allergen is not introduced into genetically modified food. The
CC present sequence represents a DNA encoding the Ara h 3 allergen. (Updated
CC on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 512 AA;
SQ
Alignment Scores:
Pred. No.: 3.64e-234 Length: 512
Score: 2652.00 Matches: 506
Percent Similarity: 99.4% Conservative: 0
Best Local Similarity: 99.4% Mismatches: 1
Query Match: 95.4% Indels: 2
DB: 2 Gaps: 1
US-10-728-323-3 (1-1524) x AAY40912 (1-512)
QY 1 CGGCAGCAACCGGAGGAGACGCTGCCAGTTCACGCGCCTCAATGCGCAGAGACCTGAC 60
Db
D 4 ArgGlnGlnProGluGluAaSnAlaLeCysGlnPheGlnArgLeuAaSnAlaGlnArgProAsp 23
QY 61 AATCGCATTTGAATCAGAGGGGGTTACATTGAGACTTGGAAACCCCAACCAACGAGGAGTTC 120
Db
D 24 AsnArgIleGluSerGluGlyGlyTyIleGluThrTrpAsnProAsnAsnGlnGluPhe 43

```

```

QY 121 GAATGCGCGCGCTGCCCTCTCTCGCTTAGCTCTCGCGCGCAACGCCCTTCGTAGGCCT 180
D 44 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 63
QY 181 TTCTACTCCAATGCTCCCGAGAGATCTTCATCCAGCAAGAGGGGATACTTTGGTTG 240
D 64 PheTySerAsnAlaProGlnGluPheIleGlnGlnGlyArgGlyTyPheGlyLeu 83
QY 241 ATATTCCTCGTTGCTCTAGACACTATGAAGACCTCACAACAAGGTGCTGCATCTCAG 300
D 84 IlePheProGlyCysProArgHisTyrgluGluProHisThrGlnGlyArgArgSerGln 103
QY 301 TCCCAAGACCAACAGACGCTCTCCAAGGAGAGACCAAGCAACAGCAACAGATAGT 360
D 104 SerGlnArgProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 123
QY 361 CACCAAGGTGCACCGTTTCGATGAGGTGTCTCATTTGCAGTTCACCGGTGTGTCT 420
D 124 HisGlnIysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143
QY 421 TTCTGGCTCTACAACGACCACGACACTGATGTTGTTGCTGTTTCTCTTACTGACACAAC 480
D 144 PheTrpLeuTyAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 163
QY 481 AACCAACGACACGACTTCATCAGTTCCCGAGAGATTCAATTTGGCTGGACACGGAG 540
D 164 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 183
QY 541 CAA-----GAGTTCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAGAGCTTA 594
D 184 GlnGluPheGluPheLeuArgTyrglnGlnSerArgGlnSerArgArgArgSerLeu 203
QY 595 CCATATAGCCCATACAGCCGCAAGTTCAGCTCAGTACAGCAAGAGAGCGTGAATTTAGCCCT 654
D 204 ProTySerProTySerProGlnSerGlnProArgGlnGluGluArgGluPheSerPro 223
QY 655 CGAGGACACACAGCCGCGAGAACGAGCAGGAGCAAGAGAAAGAAACCAAGGTGGAAC 714
D 224 ArgGlyGlnHisSerArgArgGluArgAlaGlyGlnGluGluGlnGluGlyGlyAsn 243
QY 715 ATCTTCAGCGGCTTCACGCGGAGTTCCTGGAAACAAGCCTTCCAGGTTGACGACAGACAG 774
D 244 IlePheSerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGln 263
QY 775 ATATGCAAAACCTTAGAGCGGAGACGAGAGTGAAGAGAGGAGCGCATTTGACAGTG 834
D 264 IleValGlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrVal 283
QY 835 AGCGGAGGCTCAGATCTTGACCCAGATAGAAAGAGACGTCGCCAGCAAGAGAGAGAA 894
D 284 ArgGlyGlyLeuArgIleLeuSerProAspArgGlyArgArgAlaAspGluGluGlu 303
QY 895 TACGATGAAGATGAATATGAATACGATGAAGAGATAGAGGCGTGCACGGGAAGCAGA 954
D 304 TyrAspGluAspGluTyrgluTyrgluAspGluGluAspArgArgGlyArgGlySerArg 323
QY 955 GCGAGGCGGATCGTATTGAAGAGACATCTGCACCGCAAGTGTCTAAAAGACATTGGT 1014
D 324 GlyArgGlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaIleValAsnIleGly 343
QY 1015 AGAAACAGATCCCTGACATCTACACCTCAGAGCTGTTTCACTCAAACTGCCAACGAT 1074
D 344 ArgAsnArgSerProAspIleTyrglnProGlnAlaGlySerLeuTyThrAlaAsnAsp 363
QY 1075 CTCAACCTCTTAATCTAGTGGCTTGGACCTAGTGTCTGAATATGAAATCTCTACAGG 1134
D 364 LeuAsnLeuLeuIleLeuArgTrpLeuGlyLeuSerAlaGluTyrglyAsnLeuTyrg 383
QY 1135 AATGCATTTGTTGCTCGCTCACTCAACACCAACGACACAGCATCATATATCGATTGAGG 1194
D 384 AsnAlaLeuPheValAlaHisTyrglnThrAsnAlaHisSerIleIleTyrglyLeuArg 403

```

QY 1195 GGACGGCTCAGCTGCAAGTCTGTGGACAGCAACGGCAACAGAGTGTCACGACGAGGAGCTT 1254
Db |||||
404 GlyArgAlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeu 423
QY 1255 CAAGAGGTCACGTGCTTGTGGTCCACAGAACTTCGCGTGTGGTGAAGTCCGAGAC 1314
Db |||||
424 GlnGluGlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSer 443
QY 1315 GAGAACTTCGAATACCTGGCATTCAAGACAGACTCAAGGCCAGCATAGCCAACTTCGCC 1374
Db |||||
444 GluAsnPheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAla 463
QY 1375 GGTGAAATCCCGTCATAGATAACCTGCCGAGAGGTGGTGCATAATTCATATGGCCTC 1434
Db |||||
464 GlyGluAsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeu 483
QY 1435 CAAGGGACGACGACGAGCTTAGAACACACACCCCTTCAAGTCTTCGTTCCACCG 1494
Db |||||
484 GlnArgGluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProPro 503
QY 1495 TCTCAGCAGTCTCCGAGGCTGTGGCT 1521
Db |||||
504 SerGlnGlnSerProArgAlaValAla 512

RESULT 7

AAB33601
ID AAB33601 standard; protein; 526 AA.

XX AAB33601;

XX 12-SEP-2003 (revised)

DT 22-JAN-2001 (first entry)

XX Modified Ara h 3 amino acid sequence.

XX Allergy; allergic reaction; allergen; anaphylactic antigen; peanut;
KW Ara h 1; Ara h 2; Ara h 3; epitope; binding; immunoglobulin E; IgE;
KW food antigen; sensitising; immune response; anti-allergic.

XX Arachis hypogaea.

XX WO200051647-A2.

XX 08-SEP-2000.

XX 03-MAR-2000; 2000WO-US005655.

XX 03-MAR-1999; 99US-0122960P.

PR 06-DEC-1999; 99US-00455294.

XX (MOUN) MOUNT SINAI SCHOOL MEDICINE.

XX Sampson HA;

XX WPI; 2000-611341/58.

XX Non-human animal sensitized to an antigen, useful as an animal model for
PT studying allergic reactions to allergens, such as those in food and in
PT the environment.

XX Example 6; Fig 17C; 124pp; English.

XX The present invention describes an animal model which can be used for
XX studying allergic reactions to allergens. The animal is sensitized to a
CC selected antigen by administering the antigen itself or a nucleic acid
CC encoding the antigen, where preferably the antigen is an anaphylactic
CC antigen. The sensitized animal can then be used to screen for compounds
CC which may help to prevent, ameliorate, or cure allergic conditions in
CC humans. The animal model can be used for studying allergic reactions to
CC allergens, such as those in food (peanuts, fruit, berry, nut, bean, milk,
CC dairy products), or in the environment (weed pollen, grass pollen, tree
CC pollen, mite, animal, animal dander, fungal, and insect antigens).
CC AAB33478 to AAB33601 represent sequences which are used in examples from

CC the present invention to specifically examine the peanut allergy, and the
CC peanut anaphylactic antigens Ara h 1 to 3. (Updated on 12-SEP-2003 to
CC standardise OS field)

XX Sequence 526 AA;

Alignment Scores:

7.12e-233 Length: 526
Pred. No.: 2638.00 Matches: 502
Score: 99.0% Conservative: 0
Percent Similarity: 99.0% Mismatches: 5
Best Local Similarity: 94.9% Indels: 0
Query Match: 3 Gaps: 0
DB: 3

US-10-728-323-3 (1-1524) x AAB33601 (1-526)

QY 1 CGGCAGCAACCGGAGAGAACCGGTCCAGTTCAGCGCTCAATGCCGAGAGACTGAC 60
Db |||||
5 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 24

QY 61 AATCCATTGAATCAGAGGGCGTTTACATTGAGACTTGGAAACCCCAACACAGAGAGTTC 120
Db |||||
25 AsnArgIleGluSerGluGlyTyrIleGluThrTrpAsnAlaAsnAsnGlnGluPhe 44

QY 121 GAATCGCGCGGCTCGCCCTCTCTCGCTTAGTCTCCGCCCAACGCCCTTCGTAGGCT 180
Db |||||
45 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro 64

QY 181 TTCTACTCAATGCTCCCGAGAGATCTTCAGCAGCAGGAGGGGATACCTTGGTGTG 240
Db |||||
65 PheTyrSerAsnAlaProGlnGluPheIleGlnGlnGlyArgGlyTyrPheGlyLeu 84

QY 241 ATATTCCCTGTTGTCTAGACACTATGAAGAGCTTCACACAAAGTCTGTCGATCTCAG 300
Db |||||
85 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgSerGln 104

QY 301 TCCCAAGACCAACCAAGACGCTTCAAGGAGAGAACCAAGCAACAGCAGATAGT 360
Db |||||
105 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 124

QY 361 CACGAGAGGTGCACCGTTCGATGAGGTGATCTCATTGCGATTCGCCCGGTGTGCT 420
Db |||||
125 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 144

QY 421 TTCTGGCTCTACAACGACACACTGATGTTGTTGCTGTTTCTTCTACTCACCAAC 480
Db |||||
145 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 164

QY 481 AACACGACCAACCAAGCTTTCAGTTCCTCCAGGAGATTCATTTGGCTGGGAACCGGAG 540
Db |||||
165 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 184

QY 541 CAAGAGTTCTTAAGTACCAACAAAGCAGACAAAGCAGACGAGAGCTTACCATAT 600
Db |||||
185 GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgSerLeuProTyr 204

QY 601 AGCCCATACAGCCCGGAGTTCCTGGAACAGCTTCCAGGTTGACGACAGACAGTAGTG 660
Db |||||
205 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 224

QY 661 CAGCACGCCCGCAGAGAACGACGACGACAAAGAAAGAAAGAGGTGGAAACATCTTC 720
Db |||||
225 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluAsnGluGlyGlyAsnIlePhe 244

QY 721 AGCGGCTTCACCGCGGAGTTCCTGGAACAGCTTCCAGGTTGACGACAGACAGTAGTG 780
Db |||||
245 SerGlyPheThrProGluAlaLeuPheGlnAlaPheGlnValAspAspArgGlnIleVal 264

QY 781 CAAACCTTAAGAGGGGAGACCGAGAGTGAAGAGAGGAGGCCATTTGTGACGTGAGGGA 840
Db |||||
265 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 284

QY 841 GGCCTCAGAATCTTGAGGCCCAGATAGAAAGAGACGTCGCCGACGAGAGGAATACGAT 900

Db 285 GlyLeuArgAlaLeuSerProAspAArgLysArgAlaAspGluGluGluTyrAsp 304
QY 901 GAAGATGAATATCAATACGATGAAGAGATAGAGCGTGGCAGGGAAGCAGAGCAGG 960
Db 305 GluAspGluTyrAlaTyrAspGluGluAspArgArgGlySerArgGlyArg 324
QY 961 GGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGTAAAGAACATTTGTTGAAAC 1020
Db 325 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 344
QY 1021 AGATCCCTCACATCTACACCTCAAGCTGCTCACTCAAACTGCCACGATCTCAAC 1080
Db 345 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 364
QY 1081 CTTCTAATCTAGTGGCTTGGACCTAGTCTGCAATATGCAATCTACAGGAATGCA 1140
Db 365 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 384
QY 1141 TTGTTTGTCTCCTCACTAACACACCAAGCCACAGCATCATATATCGATTGAGGGACGG 1200
Db 385 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 404
QY 1201 GCTCAGTCAAGTCGTGGACAGCAAGCGCAACAGAGTGTCACGAGGAGCTTCAAGAG 1260
Db 405 AlaHisValGlnValValAspSerAsnGlyAsnArgValIyrAspGluGluLeuGlnGlu 424
QY 1261 GGTCACTGTCTGTGGTCCACAGAACTTCGCCGTGCTGGAAGTCCACAGACGAGAAC 1320
Db 425 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 444
QY 1321 TTGCAATACGTGCATTCAAGACAGACTCAAGCCACGACATAGCCAACTTCGCGGTGAA 1380
Db 445 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 464
QY 1381 AACTCCGTATAGATAACCTGCCGAGAGGTGGTTGCAAAATTCATATGCGCTCCAAAGG 1440
Db 465 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg 484
QY 1441 GAGCAGCAAGGAGCTTAAGAACAAACACCCCTTCAAGTTCTTCGTTCCACCGTCTCAG 1500
Db 485 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProProSerGln 504
QY 1501 CAGTCTCCGAGGCTGTGGCT 1521
Db 505 GlnSerProArgAlaValAla 511

RESULT 8
AAU04711
ID AAU04711 standard; protein; 526 AA.
XX AAU04711;
XX AC
XX XX
DT 23-OCT-2001 (first entry)
XX
DE Modified anaphylactic antigen Ara h 3.
XX Ara h 3; anaphylactic antigen; immunoglobulin E; IgE; immunogenic;
KW allergy; mast cell; basophil; mouse.
XX Mus sp.
OS Synthetic.
XX WO200140264-A2.
PN
XX
XX
PD 07-JUN-2001.
XX
XX 06-DEC-2000; 2000WO-US033124.
PF
XX 06-DEC-1999; 99US-00455294.
PR 23-JUN-2000; 2000US-0213765P.
PR 27-SEP-2000; 2000US-0235797P.
XX

PA (PANA-) PANACEA PHARM LLC.
PA (OYAR-) UNIV ARKANSAS.
XX (MOUN) MOUNT SINAI SCHOOL MEDICINE NEW YORK.
PI Bannon GA, Burks WA, Caplan MJ, Sampson H, Sosin H;
XX WPI; 2001-381378/40.
XX
XX Antigenic fragments useful for reducing anaphylactic risk and reducing
PT the severity and/or number of allergic symptoms in individuals sensitive
PT to antigens, have reduced ability to bind immunoglobulin E.
XX Disclosure; Fig 12; 100pp; English.
XX

CC The sequence represents the amino acid sequence of modified anaphylactic
CC antigen Ara h 3, which has been altered to disrupt immunoglobulin E (IgE)
CC binding sites. Ara h 3 is an anaphylactic antigen (A), which was used to
CC design antigenic peptides having a reduced ability to bind IgE as
CC compared with the intact (A), or having a sequence substantially
CC identical to a portion of sequence of an antigen that includes at least
CC one IgE binding site, where at least one IgE binding site of the peptide
CC is altered. The antigenic peptides are used in a composition which is
CC useful for reducing risk or severity of allergic reaction to an antigen.
CC This is done by identifying an individual at risk of allergic reaction to
CC an antigen by identifying prior display of allergic symptoms when exposed
CC to the antigen, or a familial relationship with an individual who
CC previously displayed allergic symptoms when exposed to the antigen.
CC Following this an antigen-specific IgE present on one or more mast cells
CC or basophils in the individual's serum is identified. The individual is
CC then contacted with a peptide corresponding to a portion of the antigen,
CC which is selected, formulated, and delivered so that binding of the
CC peptide to antigen-specific IgE is reduced as compared with IgE binding
CC of intact antigen. The composition is also useful for treating and
CC preventing allergic reactions
XX
SQ Sequence 526 AA;

Alignment Scores:
Pred. No.: 7.12e-233 Length: 526
Score: 2638.00 Matches: 502
Percent Similarity: 99.0% Conservative: 0
Best Local Similarity: 99.0% Mismatches: 5
Query Match: 94.9% Indels: 0
DB: 4 Gaps: 0
US-10-728-323-3 (1-1524) x AAU04711 (1-526)

QY 1 CGGCAGCAACCGGAGGAGAACGGTCCAGTTCCAGCGCTCAATGCGCAGACCTGAC 60
Db 5 ArgGlnGlnProGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 24
QY 61 AATCGATTGAATCAGAGGGCGGTTCATTGAGACTTGGAAACCCCAACACGAGGAGTTC 120
Db 25 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnAlaAsnAsnGlnGluPhe 44
QY 121 GAATGCGCGGCGTCCGCTCTCGTTAGTCTCCGCGCAACGCCCTTCGTAGGCGCT 180
Db 45 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro 64
QY 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGGATCTTTGGGTG 240
Db 65 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 84
QY 241 ATATTCCTGGTGTCTCTAGACACTATGAAGACCTCACACAGAAGTCTCATCTCAG 300
Db 85 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgSerGln 104
QY 301 TCCCAAGACCCACCAAGAGCTCTCCAGGAGAGACCAAGCCAAACAGACAGATAGT 360
Db 105 SerGlnArgProArgArgLeuGlnGlyGluAspGlnSerGlnGlnGlnArgAspSer 124
QY 361 CACCAGAAGGTGCACCGCTTTCGATGAGGGTGTCTCATTCAGTTCACCGGTGTGCT 420

Db 125 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 144
QY 421 TTCTGGCTCTACAACACACACACACTGATGTTGTTGTTCTTCTTACTGACACCAAC 480
Db 145 PheTrpLeuTyrrAsnAspHisAspThrAspValAlaValSerLeuThrAspThrAsn 164
QY 481 AACACGACAAACAGCTTGATCAGTTCCTCCAGGAGATTCAATTTGGCTGGGAACGCGAG 540
Db 165 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 184
QY 541 CRAGAGTTCTTAGGTTACGACGACAAAGCAGACAAAGCAGACGACGAGAGCTTACCATAT 600
Db 185 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 204
QY 601 AGCCCATACAGCCCGCAAGTTCAGCTAGACAAAGAGAGCGTAATTTAGCCCTCGAGA 660
Db 205 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgPheSerProArgGly 224
QY 661 CAGCACAGCCGACAGAAACGAGCAGACAAAGAAAGAAACGAAAGTGGAAACATCTTC 720
Db 225 GlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIlePhe 244
QY 721 AGCGGCTTCACGCCGAGTTCCTGGAAACAAGCTTCAGGTTGACGACAGACAGATAGTG 780
Db 245 SerGlyPheThrProGluAlaLeuPheGlnAlaPheGlnValAspAspArgGlnIleVal 264
QY 781 CAAACCTAAGAGCCGAGACCGAGAGTGAAGAGAGGAGCCATTGTGACAGTGAAGGGA 840
Db 265 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 284
QY 841 GGCCTCAGAAATCTTGAGCCCGATAGAAAGAGACGTCGCCGACGAAAGAGGAATACGAT 900
Db 285 GlyLeuArgAlaLeuSerProAspArgLysArgAlaAspGluGluGluTyrAsp 304
QY 901 GAAGATGAATATCAATACATGAAGAGGATAGAGCGTCGCGGGAAGCAGAGCAGG 960
Db 305 GluAspGluTyrrAlaTyrrAspGluGluAspArgArgGlyArgGlySerArgGlyArg 324
QY 961 GGGATGGTATTGAAGAGACGATCTGCACCGCAAGTCTAAAGAAACATTCGTGAAGAC 1020
Db 325 GlyAsnGlyLeuGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 344
QY 1021 AGATCCCTCGACATCTACAAACCTCAAGCTGGTTCTCAAACTGCCCAACGATCTCAAC 1080
Db 345 ArgSerProAspIleTyrrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 364
QY 1081 CTTCTAATACTTAGTGGCTTGGACCTAGTGTGTAATATGGAATCTTACAGGATGCA 1140
Db 365 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrrArgAsnAla 384
QY 1141 TTGTTTGTGCTCCTACACACCAACGACACGACATCATATATCGATTGAGGGGACGG 1200
Db 385 LeuPheValAlaHisTyrrAsnThrAsnAlaHisSerIleIleTyrrArgLeuArgGlyArg 404
QY 1201 GTCACGTCGAAGTCTGGACAGCAACGCGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
Db 405 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrrAspGluGluLeuGlnGlu 424
QY 1261 GGTACGTCGTGTGTGTCACAGAACTTCGCGCTCGTGGAAAGTCCAGAGCGAGAAC 1320
Db 425 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 444
QY 1321 TTCGAATAGTGCATTCAAGACAGACTCAAGCCCGACGATAGCCCAACCTCCCGGTGAA 1380
Db 445 PheGluTyrrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 464
QY 1381 AACTCCGTCATAGATAACCTCGCGGAGAGGTGGTTGCAAAATTCATATGCGCTTCCAAAGG 1440
Db 465 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrrGlyLeuGlnArg 484
QY 1441 GAGCAGGACGAGGAGTCTAAGAACACAAACCCCTTCAAGTCTTCTGTTCCACCGCTCAG 1500
Db 485 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProProSerGln 504

QY 1501 CAGTCTCCGAGGGCTGTGGCT 1521
Db 505 GlnSerProArgAlaValAla 511

RESULT 9

AAU05036
ID AAU05036 standard; protein; 526 AA.

AC AAU05036;

XX 24-OCT-2001 (first entry)

XX Modified anaphylactic peanut antigen Ara h 3.

XX Anaphylactic antigen; Ara h 3; peanut; desensitisation; antigen; allergy;
immunotherapy; rush immunotherapy; anaphylaxis; asthma; food allergen;
milk allergen; shellfish allergen; latex; drug; environmental allergen;
grass pollen; ovalbumin; insect venom; peanut oil; immunoglobulin E; IgE.
XX Arachis sp.
OS Synthetic.

XX WO200139799-A2.

PN 07-JUN-2001.

XX 06-DEC-2000; 2000WO-US033125.

XX 06-DEC-1999; 98US-00455294.

PR 23-JUN-2000; 2000US-0213765P.

PR 27-SEP-2000; 2000US-0235797P.

XX (PANA-) PANACEA PHARM LLC.

XX Caplan M;

XX WPI; 2001-408258/43.

XX Preventing allergic response against antigens, e.g. food and
environmental allergens such as peanut allergen or grass pollen, latex or
drug, comprises administering agent, e.g. a peptide, that blocks antigen
binding sites on offending IgE.

XX Disclosure; Fig 2; 76pp; English.

XX The sequence represents the amino acid sequence of modified anaphylactic
peanut antigen Ara h 3. The protein is used for active or passive
desensitisation of an individual to an antigen; for alleviating or
preventing allergic reactions and for decreasing the risk of allergic
reactions during immunotherapy or rush immunotherapy, anaphylaxis and
asthma. The antigen may be a food allergen (e.g. peanut or milk
allergen), shellfish allergen, environmental allergen (e.g. grass pollen
or tree pollen), latex, drug, pollen, ovalbumin, an insect venom antigen
or predominantly linear epitopes. The protein is useful for protecting an
individual against subsequent inadvertent or intentional exposure to
antigen, e.g. receiving blocking agent before eating a chocolate bar
which may inadvertently contain peanut components or before eating foods
prepared using peanut oil. Administration of the blocking agents does not
result in cross-linking of anti-antigenic immunoglobulin E (IgE). After
exposure to the agent, the individual's antigen sensitivity is at least
temporarily reduced. Only those IgE molecules that bind the offending
antigen and contribute to the risk of an allergic response are blocked

XX Sequence 526 AA;

Alignment Scores:

Pred. No.: 7,12e-233 Length: 526
Score: 2638.00 Matches: 502
Percent Similarity: 99.0% Conservative: 0
Best Local Similarity: 99.0% Mismatches: 5
Query Match: 94.9% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-3 (1-1524) X AAU05036 (1-526)

Qy	1	CGG	CAG	CAACCGG	GAGGAACGG	TGTC	CGCAGTTC	CAGCGCTCAAT	CGCGCAGACCT	GAC	60
Db	5	Arg	Gln	Pro	Glu	Glu	Asn	Ala	Cys	Gln	24
Qy	61	AAT	CGCATTC	CAATC	CAGAGGCGG	TTAC	ATTGAGACTT	TGGAAACCC	CAACACCA	CAGGAGTTC	120
Db	25	Asn	Arg	Ile	Glu	Arg	Gly	Gly	Tyr	Ile	44
Qy	121	GAAT	CGCCCGG	CGTGC	CGCTCT	CTCGT	TAGTTC	CTCGCGCAAC	CGCCCTTC	TGTAGGCCT	180
Db	45	Glu	Cys	Ala	Gly	Val	Ala	Leu	Ser	Arg	64
Qy	181	TTT	TACTT	CCAAT	GTCTCC	CCAGGAGAT	CTT	CAT	TCCAGCAAG	AGGGGAT	240
Db	65	Phe	Tyr	Ser	Asn	Ala	Pro	Gln	Glu	Ile	84
Qy	241	ATAT	TCCCTCG	TGTCCT	TAGACACT	ATG	AAAGAGCCT	CACACACA	AGGTGCT	GCATCT	300
Db	85	Ile	Phe	Pro	Gly	Cys	Pro	Arg	His	Tyr	104
Qy	301	TCCC	CAAAAGAC	CAAC	AAGACG	TCTCCA	AGGAGAAG	CAACCAAA	CGCAAC	GACGATAGT	360
Db	105	Ser	Gln	Arg	Pro	Pro	Arg	Leu	Gln	Gly	124
Qy	361	CAC	CAGAGGT	GCACCG	TTTC	GTATG	AGGGTGAT	CTCAT	TCCAGTTC	CCACCGTGT	420
Db	125	His	Gln	lys	Val	His	Arg	Phe	Asp	Glu	144
Qy	421	TTT	CGGCTT	ACACG	CCACG	ACAC	TGATCT	TGCTGCT	GTCTCT	TACTGAC	480
Db	145	Phe	Trp	Leu	Tyr	Asn	Asp	His	Asp	Thr	164
Qy	481	AAC	CAACGAC	CAAC	CACTTG	ATCAGT	TCCCCAGG	AGATTCA	ATT	TGGCTGG	540
Db	165	Asn	Asn	Asp	Asn	Gln	Leu	Asp	Gln	Phe	184
Qy	541	CAAG	ATGTTT	TAAGG	TACC	GACCA	CAAA	AGCAGACA	AAAGCAGACA	GAAGAACT	600
Db	185	Gln	Glu	Phe	Leu	Arg	Tyr	Gln	Gln	Gln	204
Qy	601	AGCC	ATACAC	CCCG	CAAA	GTCC	GTAGACA	CAAGAAG	CGTGA	ATT	660
Db	205	Ser	Pro	Tyr	Ser	Pro	Gln	Ser	Gln	Pro	224
Qy	661	CAG	CACAGCC	CGACAG	CAACG	ACGAGCA	GAAGA	AAAAACGA	AGTGG	AAAAATCT	720
Db	225	Gln	His	Ser	Arg	Arg	Glu	Arg	Ala	Gly	244
Qy	721	AGC	GGCTTC	CACCG	CGGAGT	TCTCG	GAACAAC	AGCTTC	CGAGT	TGACAGAT	780
Db	245	Ser	Gly	Phe	Thr	Pro	Glu	Ala	Leu	Phe	264
Qy	781	CAAA	ACCTTA	CAGCG	GACCG	ACGAG	GTCAACA	GAGGGAG	CCATTG	TGACAGT	840
Db	265	Gln	Asn	Leu	Arg	Gly	Glu	Thr	Glu	Ala	284
Qy	841	GGC	CTCAGA	ATCTT	GAGCC	CAATAG	ATAA	GAGAG	CGTGC	CGACGA	900
Db	285	Gly	Leu	Arg	Ala	Leu	Ser	Pro	Asp	Arg	304
Qy	901	GAA	GATG	GAAT	TATAC	GATGA	AGAGAT	TAGA	GGGTGC	CAGGGA	960
Db	305	Glu	Asp	Glu	Tyr	Ala	Tyr	Asp	Glu	Ala	324
Qy	961	GGG	AATG	TAT	TGA	AGAC	ACCAT	CTGC	ACCCCA	AGTGTCT	1020
Db	325	Gly	Asn	Gly	Ile	Glu	Thr	Ile	Cys	Thr	344
Qy	1021	AGAT	CCCC	TGAC	TCTCA	CAAC	CCCTCA	AGCTGG	TTCAC	TCAAAA	1080

Db	345	ArgSerProhspIleTyrAsnProGlnAlaGlySerLeuIysThrAlaAsnAspLeuAsn	364
Qy	1081	CTTCTAATACTAGGTGGCTTGGACCTAGTGTGGAATATGGAATCTCTACAGGAATGCA	1140
Db	365	LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla	384
Qy	1141	TTGTTTGTGCTCACTACAACACCAACGACACAGCATCATATATCGATTGAGGGACGG	1200
Db	385	LeuPheValAlaAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg	404
Qy	1201	GCTCAGTGCACAGTCGTGGACAGCAACGCGCAACAGAGTGTACGACGAGGAGCTTCAAGAG	1260
Db	405	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu	424
Qy	1261	GGTCAGTGTCTGTGTGGCCACAGAACTTCGCCGTCGCTGGAAAAGTCCCAAGACGAGAAC	1320
Db	425	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn	444
Qy	1321	TTCAATPACGTGGATTCAAGACAGACTCAAGGCCCGCCAGCATAGCCAACTTCGCGCGGTGAA	1380
Db	445	PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	464
Qy	1381	AACTCCGTCATAGATAACTGTCGCGGAGGAGTGGTTCGCAAAATTCATATGCGCTCCAAAGG	1440
Db	465	AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg	484
Qy	1441	GAGCAGCAAGGACGCTTAAGAACAAACACCCCTTCAAAGTTCCTTCGTTCCACCGTCTCAG	1500
Db	485	GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheValProProSerGln	504
Qy	1501	CAGTCTCCGAGGCTGTGGCT	1521
Db	505	GlnSerProArgAlaValAla	511
RESULT 10			
ADO38357			
XX	ID	ADO38357	standard; protein; 507 AA.
XX	AC	AC	
XX	AC	AC	
XX	AC	AC	
DT	15-JUL-2004	(first entry)	
XX	Peanut allergen Ara h 3.		
XX	Antibacterial; Virucide; Fungicide; Antiparasitic; Antiarthritic;		
KW	Antirheumatic; Neuroprotective; Antiinflammatory; Dermatological;		
KW	Immunosuppressive; Antidiabetic; Antithyroid; Aniasthmatic;		
KW	Antiallergic; Cytostatic; Antiporiatic; Gene Therapy; Vaccine;		
KW	MHC Class II; II-key motif; Immune response; Anthrax; EBOLA; HIV;		
KW	influenza; vaccinia virus; infection; bacterium; virus; parasite; fungus;		
KW	ricketsia; rheumatoid arthritis; multiple sclerosis;		
KW	lupus erythematosus; diabetes mellitus; myasthenia gravis;		
KW	autoimmune thyroiditis; scleroderma; dermatomyositis; pemphigus; asthma;		
KW	allergic rhinitis; topical dermatitis; colitis; cancer; psoriasis;		
KW	adenoma; peanut; Ara h 3.		
XX	Arachis hypogaea.		
OS	US2004058881-A1.		
XX	25-MAR-2004.		
XX	24-SEP-2002; 2002US-00253286.		
XX	24-SEP-2002; 2002US-00253286.		
PR	(ANTI-) ANTIGEN EXPRESS INC.		
XX	Humphreys RE, Xu M;		
PI	WPI; 2004-294259/27.		
XX	New non-naturally occurring protein or polypeptide modified by		
PT			

PT recombinant DNA techniques, useful for treating multiple sclerosis,
 PT diabetes mellitus, myasthenia gravis, scleroderma, allergic rhinitis,
 PT colitis, cancer or psoriasis.

XX Example 3; Page 21-22; 90pp; English.

XX The invention relates to a non-naturally occurring protein or polypeptide
 CC (i) modified by recombinant DNA techniques comprising: a C-terminal
 CC element comprising an MHC Class II-presented epitope; an N-terminal
 CC element comprising an Ii-key motif; and an intervening element comprising
 CC a sequence of 4-11 amino acid residues where the modification by
 CC recombinant DNA techniques taking place within elements (b) and (c). Also
 CC described are methods for: suppressing or enhancing an immune response
 CC directed toward an MHC (major histocompatibility complex) Class II-
 CC presented epitope of interest. Suppressing an immune response directed
 CC toward an MHC Class II-presented epitope of interest comprises: providing
 CC a nucleic acid sequence encoding the MHC Class II-presented epitope of
 CC interest; the nucleic acid sequence encoding an Ii-key motif located 4-11
 CC amino acids upstream from the N-terminal residue of the MHC Class II-
 CC presented epitope of interest; and modifying the Ii-key motif to decrease
 CC its conformance to the archetypal Ii-key regulatory motif. Enhancing an
 CC immune response directed toward an MHC Class II-presented epitope of
 CC interest comprises: providing a nucleic acid sequence encoding the MHC
 CC Class II-presented epitope of interest; the nucleic acid sequence lacking
 CC an Ii-key motif located 4-11 amino acids upstream from the N-terminal
 CC residue of the MHC Class II-presented epitope of interest; and modifying
 CC the nucleic acid sequence to introduce an Ii-key motif appropriately
 CC spaced from the MHC Class II-presented epitope. The protein or
 CC polypeptide of interest corresponds to a protein or polypeptide encoded
 CC by an infectious pathogen selected from anthrax, EBOLA, HIV or influenza,
 CC preferably vaccinia virus. The non-naturally occurring protein or
 CC polypeptide (I) modified by recombinant DNA techniques is useful for
 CC treating infectious diseases caused or associated with infection by a
 CC bacterium, virus, parasite, fungus, rickettsia or other infectious
 CC agents. It is also useful for treating rheumatoid arthritis, multiple
 CC sclerosis, lupus erythematosus, diabetes mellitus, myasthenia gravis,
 CC autoimmune thyroiditis, scleroderma, dermatomyositis, pemphigus, asthma,
 CC allergic rhinitis, topical dermatitis, colitis, cancer, psoriasis or
 CC adenomas. The present sequence represents the amino acid sequence of
 CC peanut allergen Ara h 3 used in the invention.

XX Sequence 507 AA;

Alignment Scores:
 Pred. No.: 9,08e-231 Length: 507
 Score: 2615.00 Matches: 498
 Percent Similarity: 98.2% Conservative: 0
 Best Local Similarity: 98.2% Mismatches: 9
 Query Match: 94.1% Indels: 0
 DB: 8 Gaps: 0

US-10-728-323-3 (1-1524) x ADO38357 (1-507)

QY 1 CGGACCAACCGAGAGAACGGTCCAGTTCACGGCTCAATCGCGAGACCTGAC 60
 Db 1 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlyArgLeuAsnAlaGlnArgProAsp 20
 QY 61 AATCGCATTAATCAGACGGCGTTACATTGAGACTTGCAACCCCAACACGAGGATTTC 120
 Db 21 AsnArgileGluSerGluGlyGlyfyrileGluThrTrpAsnProAsnAsnGlyGluPhe 40
 QY 121 GAATGCGCGGGGTCCCTCTCTCGCTTAGTCTCCGCGCAACGGCCCTTCGTAGGCCT 180
 Db 41 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro 60
 QY 181 TTCTACTCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGGGATCTTGGTTG 240
 Db 61 PheTyrSerAsnAlaProGlnGluIlePheIleGlyGlyArgGlyTyrPheGlyLeu 80
 QY 241 ATATTCCCTGGTGTCTAGACACTATGAGAGCTTCACACACAGGTCGTGATCTCAG 300
 Db 81 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlyGlyArgArgSerGln 100

QY 301 TCCCAAAGACCAACCAAGACGCTCTCCAAGGAGAAGACCAAGCCAAACAGACAGATAGT 360
 Db 101 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 120
 QY 361 CACGAGAAGGTGCACCGTTTCGATGAGGGTGATCTCATTTGCTTCCCAACCGGTGTGCT 420
 Db 121 HisGlnLysValHisArgPheAspGluGlyPheLeuIleAlaValProThrGlyValAla 140
 QY 421 TTCTCGCTCTACAACGACACGACACTCATGTGTGTGTGTGTCTTCTTACTCACCAAC 480
 Db 141 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 160
 QY 481 AACACGACCAACAGCTTGTATGATTTCCCGAGAGATTCATTTGGCTGGGAACCGAG 540
 Db 161 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 180
 QY 541 CNAAGATTCTTAAGTACAGCAACAAAGCAGACAAAGCAGACGACGAGAGCTTACCATAT 600
 Db 181 GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgGlnSerArgSerLeuProTyr 200
 QY 601 AGCCCATACAGCCCGCCAAAGTCAGCTTAGACAAGAGCGGTGAATTTAGCCCTCGAGA 660
 Db 201 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 220
 QY 661 CAGCACACCCCGACAGAACGACGAGCAAGAACAAAGAACAAAGAGTGGAAACATCTTC 720
 Db 221 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluAsnGluGlyGlyAsnIlePhe 240
 QY 721 AGCGGCTTCACGCGGAGTTCTCTGAAACAGGCTTCCAGGTTGACGACAGACAGTAGTG 780
 Db 241 SerGlyPheThrProGluPheLeuGlnAlaPheGlnValAspAspArgGlnIleVal 260
 QY 781 CAAACACCTTAAGAGCGGACACCGAGAGTGAAGAGAGGAGGACCTTGTGACAGTGAGGGA 840
 Db 261 GlnAsnLeuArgGlyGlyLutThrSerGluSerGluGluGlyAlaIleValThrValArgGly 280
 QY 841 GGCCTCAAAATCTTGAGCCCGACATAGAAGAGACCTGCGCAGAAAGAGAGAGAAATACGAT 900
 Db 281 GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrAsp 300
 QY 901 GAAGATGAATATGATACGATGAGAGATAGAACGCTGCGCAGGGGAGCAGGAGCGAGG 960
 Db 301 GluAspGluTyrGlyTyrAspGluGluAspArgArgArgGlyArgGlySerArgGlyArg 320
 QY 961 GGGAAATGCTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAGAAACAACTTGGTGAAC 1020
 Db 321 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysLysAsnIleGlyArgAsn 340
 QY 1021 AGATCCCTGACATCTACAACCCCTCAAGCTGTTCACCTCAAAACCTGCCAACCATCTCAAC 1080
 Db 341 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 360
 QY 1081 CTTCTAATACTTAGTGTGCTGGACCTAGTCTGAATATGGAATCTCTACAGGAATGCA 1140
 Db 361 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 380
 QY 1141 TTGTTGTGCTGCTACTACAACCAACCAACGACACGACATCATATATCGATTGAGGGGACGG 1200
 Db 381 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 400
 QY 1201 GCTCAGCTGCAAGTCGTGGACAGCAACCGCAACAGAGTGTACGACGAGAGGCTTCAAGAG 1260
 Db 401 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 420
 QY 1261 GGTACGCTGCTGTGGTCCACAGAACTTCGCCGCTCGCTGGAAAGTCCGAGAGCCGAGAC 1320
 Db 421 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 440
 QY 1321 TTCGAATACGTGGCATTCACAGACACTCAAGGCCCGACGATAGCCAACTCCGCCGTGAA 1380
 Db 441 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 460
 QY 1381 AACTCCGTATAGATAACCTCGCGGAGGAGGTGTTGCAATTCATATGCGCCTCCAAAGG 1440


```
QY 1252 CTTCAAGAGGTCACGTGCTTGTGTGGTCCACAGAACTTCGCCGTCGTGGAAGTCCCAAG 1311
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 392 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1312 AGCAGAACTTCAATACGTGGATTCAAGACAGACTCAAGGCCAGCATAGCCACCTC 1371
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 412 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431
QY 1372 GCGGGTGAACCTCCGTCATAGATAACCTGCCGGAGGAGTGGTTCAAATTTCATATGGC 1431
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 432 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 451
QY 1432 CTCCTCAAGGAGGAGCAGCAAGCGCTTAAGAAACAACACCCCTTCAAGTTCCTCGTTCCA 1491
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 452 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 471
QY 1492 CCGTCTCAGCAGTCTCGAGGCGTGTGGCT 1521
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 472 ProLysGluSerGlnArgValValAla 481
RESULT 12
ADH89253
ID ADH89253 standard; protein; 481 AA.
XX
XX ADH89253;
XX
XX 06-MAY-2004 (first entry)
XX
XX G. max glycinein G3 subunit.
XX
XX double stranded RNA; storage protein; 2S-albumen; 7S-globulin;
XX 11S/12S-globulin; zein-prolamine; homogenistate metabolic pathway;
XX pharmaceutical; plant; abiotic stress; fatty acid composition;
XX lipid composition; oil composition; carbohydrate composition; colour;
XX pigmentation; pathogen resistance; fruit ripening delay; ageing;
XX male sterility; lignin; fibre; cotton; Vitamin E synthesis; nicotine;
XX caffeine; theophylline; threonine biosynthesis; glycinein.
XX
XX Glycine max.
XX
XX WO2003078629-A1.
XX
XX 25-SEP-2003.
XX
XX 17-MAR-2003; 2003WO-EP002735.
XX
XX 20-MAR-2002; 2002DE-01012892.
XX
XX (BADI ) BASF PLANT SCI GMBH.
XX
XX Kock M, Bauer J;
XX
XX WPI; 2003-803889/75.
XX
XX N-PSDB; ADH89252.
XX
XX Reducing expression of at least two target genes, useful e.g. for
XX producing transgenic plants, using partly double-stranded interfering
XX RNA.
XX
XX Disclosure; SEQ ID NO 28; 228pp; German.
XX
XX This invention describes a novel method for reducing the expression of at
XX least two different endogenous target genes in a eukaryotic cell or
XX organism by introducing an RNA molecule that is at least partly double
XX stranded. The transcribed RNAs from at least two target genes have
XX homology below 90% and the RNA molecule is formed as a single, self-
XX complementary molecule. At least one of the double-stranded structures
XX formed from individual sense sequences has an even number of repeats of
XX 21 or 22 bp. The RNA molecule may include an intron-encoding sequence. At
XX least two target genes are selected from different classes of storage
XX protein genes, i.e. 2S-albumen, 7S- or 11S/12S-globulins or zein-
XX prolamine and at least one of the sense sequences is identical to storage
```

```
CC protein sequences or genes in the homogenistate metabolic pathway or
CC enzyme types, e.g. acetyl transferases, thioesterases, (de)branching
CC enzymes or cellulases. The RNA of the invention, also related cassettes,
CC expression systems, vectors and transgenic organisms are used for
CC preparation of pharmaceuticals, in biotechnological processes and plant
CC biotechnology, specifically in plants to improve protection against
CC abiotic stress, to modify composition and/or content of fatty acids,
CC lipids and oils, to modify carbohydrate composition, to alter colour or
CC pigmentation, to reduce content of storage proteins, to increase
CC resistance to pathogens, to inhibit stem break, to delay fruit ripening
CC or aging, to induce male sterility, to reduce content of toxic or
CC unwanted components, to modify lignification and/or lignin content, to
CC modify the fibre component in foods or fibre quality in cotton, to reduce
CC susceptibility to shock, to increase synthesis of vitamin E, to reduce
CC contents of nicotine, caffeine or theophylline and to increase methionine
CC content, by reducing threonine biosynthesis. The method provides a rapid
CC and efficient way of reducing gene expression, can inhibit more than one
CC target gene, prevents development of multiple phenotypes (since the
CC transcription rate is the same for all RNA sequences, significantly
CC reducing the selection process required to produce an organism with
CC effective suppression of all target genes), avoids problems of epigenic
CC gene silencing, does not require synthesis of individual RNA sequences
CC and the method can be applied to plants with complex (polyploid) genomes.
CC No interference between the individual RNA sequences occur. This sequence
CC represents a protein encoded by a target gene used in the method of the
XX invention.
XX
XX SQ Sequence 481 AA;
XX
XX Alignment Scores:
XX Pred. No.: 1.5e-128 Length: 481
XX Score: 1502.00 Matches: 297
XX Percent Similarity: 71.6% Conservative: 68
XX Best Local Similarity: 58.2% Mismatches: 91
XX Query Match: 54.0% Indels: 54
XX DB: 7 Gaps: 7
XX
XX US-10-728-323-3 (1-1524) x ADH89253 (1-481)
QY 1 CGSCAGCAACCGGAGGAGACGCGTCCAGCGCTCAATGCGCAGAGACCTGAC 60
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 23 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuIlePheAsp 42
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 AATCGCATTGAATCAGAGGCGGTTCATTGAGACTTGAACCCCAACACAGGAGTTC 120
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 43 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 62
QY 121 GAATGCGCGCGGTGCGCTCTCTCGTTAGTCTCCCGCCGCAACGCCCTTCGTAGGCCT 180
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 82
QY 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAGGAGGAGGATCTTGGGTG 240
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 83 SerTyrThrAsnAlaProGlnGlnIleTyrIleGlnGlnGlySerGlyIleGlyMet 102
QY 241 ATATTCCCTGGTGTCTAGACACTATGAAGAGCTTCACACACAGGTCGTGATCTCAG 300
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 119
QY 301 TCCCAAGACCAACCAAGAGCGTCTCAAGGAGAGAACCAAGCCACAGCAGAGATAGT 360
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 120 SerSerArgPro-----GlnAspArg 126
QY 361 CACCAAGAGGTCACCGGTTTCATGAGGATCTCATTCAGTTCAGTTCACCGGTGTGCT 420
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 127 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146
QY 421 TTCTGGCTCTACAAGCAGCACCACACTGATGTTGTGTGTTCTTCTTCTTACTCACACCAAC 480
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 147 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166
QY 481 AACACGACCAACCGCTTGATCAGTTCCTCCAGGAGATTCATTTGGCTGGGAACCGAG 540
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 167 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
QY 541 CRAAGAGTTCTTAAGGTACCAAGCAACAAAGACAGACAGCAAGAGACTTTACCATAT 600
Db 187 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly----- 199
QY 601 AGCCCATACAGCCGCAAGTACGCTAGCAACAAGAGCGGTGAATTTAGCCCTCAGGA 660
Db 200 -----Gly 200
QY 661 CAGCACAGCCGACAGAGAACGAGCAGCAGCAAGAGAAAGAAACGAAGTGGAAACATCTTC 720
Db 201 ThrGlnSerGlnLysGlyLysArgGlnGlnGlnGluGlnGluGlnGlyGlySerIleLeu 220
QY 721 AGCGGCTTCACGCCGAGTTCTCTGGAACAAGCCCTTCAGGTTGACACAGACAGATAGTG 780
Db 221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239
QY 781 CAAACCTAAGAGCGCAGCAGCAGAGAGTGAAGAGAGGAGCCATTGTGACAGTGGAGGGA 840
Db 240 ArgLysLeuGlnGlyGluAsnGluGluGluGlyAlaIleValThrValLysGly 259
QY 841 GGCTCAGATCTTGAGCCCA-----GATAGAAAGAGACGTGCCGACGAAGAAGAG 891
Db 260 GlyLeuSerValIleSerProProThrGluGlnGlnGlnArgProGluGluGluGlu 279
QY 892 GAATACGATGAAGATGAATATGAATAGATGAAGAGGATAGAAGCGGTGSCAGGGGAAGC 951
Db 280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294
QY 952 AGAGCGAGGGGATGTTATGAAGACAGCATCTGCACCCCAAGTGTAAAGACATTT 1011
Db 295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311
QY 1012 GGTAGAAACAGATCCCTCGACATCTACAACCTCAAGCTGGTTCACCTCAAAACTGCCAAC 1071
Db 312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrAlaThr 331
QY 1072 GATCTCAACCTTCTAATCTAGTGGCTTGGACCTAGTCTGAATATGAAATCTCTAC 1131
Db 332 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 351
QY 1132 AGGAATGATTTGTTCGCTCACTACACACCAACGACACACATCATATCATGTTG 1191
Db 352 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371
QY 1192 AGGGACGGCTCACGTGCAAGTCGTGCAGCAGCAACGGCAACAGAGTGTACGACGAGAG 1251
Db 372 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 391
QY 1252 CTTCAAGAGGTCACGTGCTTGTGTGCCACAGAACTTCGCGTGCCTGGAAATGCCAG 1311
Db 392 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411
QY 1312 AGCAGAACTTCGAATACGTGGCATTCAGACAGACTCAAGGCCCATAGCCCACTC 1371
Db 412 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431
QY 1372 GCGGTCGAAACTCCGTCATAGATAACTTCGCGGAGGAGGTGGTTGCAAAATTCATATGGC 1431
Db 432 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 451
QY 1432 CTCGAAAGGAGCAGGACGAGCTTAAGAACAACACCCCTTCAAGTTCTTCGTTCGA 1491
Db 452 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 471
QY 1492 CCGTCTCAGAGCTCCGAGGGCTCGCT 1521
Db 472 ProLysGluSerGlnArgValValAla 481
RESULT 13
ADL90187
ID ADL90187 standard; protein; 481 AA.

XX ADL90187;
AC 20-MAY-2004 (first entry)
XX Soybean glycinin G3 protein.
XX immunomodulator; immunotherapy; allergen characterisation;
KW immunoglobulin E; allergen sensitivity; soybean; glycinin G3;
KW acidic protein.
XX Glycine max.
XX US2003166518-A1.
XX 04-SEP-2003.
XX 12-JAN-2001; 2001US-00759967.
XX 13-JAN-2000; 2000US-0175948P.
PR 03-MAR-2000; 2000US-0186724P.
XX (BEAR/) BEARDSLEE T A.
PA (ZEEC/) ZEECE M G.
PA (SARA/) SARATH G.
PA (MARK/) MARKWELL J P.
XX Beardslee TA, Zeece MG, Sarath G, Markwell JP;
PI WPI; 2003-898094/82.
XX Allergen characterization comprises obtaining a recombinant fusion
PT protein and detecting the binding of immunoglobulin E molecules in the
PT biological sample to the recombinant fusion protein.
XX Disclosure; SEQ ID NO 21; 34pp; English.
XX The invention describes a method of allergen characterisation comprising:
CC obtaining a recombinant fusion protein; attaching the recombinant fusion
CC protein to a substrate through the native protein; contacting the
CC recombinant fusion protein attached to the substrate with a biological
CC sample from an individual; and detecting the binding of immunoglobulin E
CC molecules in the biological sample to the recombinant fusion protein.
CC Also described are: a method for determining the sensitivity of an
CC individual to a suspected allergen; a method for determining the amount
CC of immunoglobulin E specific for an allergen in a biological sample; a
CC method of immunotherapy; a method of allergen characterisation; a method
CC for determining the sensitivity of an individual to a suspected allergen;
CC a method of determining the amount of immunoglobulin E specific for an
CC allergen in a biological sample; a kit comprising the recombinant fusion
CC protein and instructions for using the recombinant fusion protein to
CC determine IgE binding to the know or suspected allergen; and a method for
CC epitope determination. The method is useful for characterising allergens.
CC This is the amino acid sequence of soybean glycinin G2 acidic protein
CC that can be used to demonstrate the methods of the invention.
XX SQ Sequence 481 AA;
Alignment Scores:
Pred. No.: 1.5e-128 Length: 481
Score: 1502.00 Matches: 297
Percent Similarity: 71.6% Conservative: 68
Best Local Similarity: 58.2% Mismatches: 91
Query Match: 54.0% Indels: 54
DB: 7 Gaps: 7
US-10-728-323-3 (1-1524) x ADL90187 (1-481)
QY 1 CGGCAGCAACCGGAGGAGAACGGCTGCCAGTTCAGGCGCTCAATGCGCAGACCTGAC 60
Db 23 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 42
QY 61 AATCGCATTGAATCAGAGGGCGGTACATTGAGACTTGGAAACCCCAACACAGGAGTTC 120

storage protein genes. Most preferably a plant cell is stably transfected with a recombinant expression construct, then regenerated to plants that express the incorporated sequence. The expression constructs particularly contain a seed-specific promoter and they are introduced into plants by standard methods, e.g. via *Agrobacterium*. The preferred storage proteins of the invention are 2S-albumins, 7S or 11S/12S-globulins or zein-prolamines. Transgenic organisms produced by the new method are used for production of oils, fats, free fatty acids or their derivatives, useful as foods, animal feeds, pharmaceuticals and fine chemicals. This sequence represents a storage protein used to illustrate the method of the invention.

[illegible]


```

||||| 399 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaIalaArgSerGln 418
||||| 1312 AGCGAGAACTTCGAATACGTGGCATTCAAGACACAGACTCAAGGCCAGCATAGCCAAACCTC 1371
||||| 419 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 438
||||| 1372 GCCGGTGAAAACTCCGTCATAGATAACCTGCCGAGAGGTGGTTGCAAAATTCATATGGC 1431
||||| 439 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 458
||||| 1432 CTCCAAAGGAGCAGCAGGCAGCTTAAGAACAAACCCCTTCAAGTTCTTCGTTCCA 1491
||||| 459 LeuArgGlnGlnAlaArgGlnValLysAsnAsnProPheSerPheLeuValPro 478
||||| 1492 CCGTCTCAGCAGTCTCCGAGGGCTGTGGCT 1521
||||| 479 ProLysGluSerGlnArgArgValAla 488
```

Search completed: May 15, 2006, 21:45:05
Job time : 259.261 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 22:18:32 ; Search time 9.16611 Seconds
(without alignments)
4123.811 Million cell updates/sec

Title: US-10-728-323-3

Perfect score: 2779

Sequence: 1 cggcgagcaccggaggagaa.....ctccgagggtgtggtctaa 1524

Scoring table: BLOSUM62

Xgapcp 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DRV=xlp
-Q=/abss/ABSSWEB_spool/US10728323/runat_15052006_172137_22463/app_query.fasta_1
-DB=Issued Patents AA -OFFMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs02p -USER=US10728323 @C@N_1_131 @runat_15052006_172137_22463
-NCPU=6 -ICPU=3 -NO MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5 COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1466	52.8	495	2	US-09-805-694B-16
2	811	29.2	492	2	US-09-462-720-2
3	623.5	22.4	483	2	US-10-053-410-6
4	293	10.5	141	2	US-09-645-593-12
5	227.5	8.2	85	2	US-09-645-593-10
6	215	7.7	165	2	US-09-645-593-11
7	184.5	6.6	489	2	US-09-424-283-3
8	173.5	6.2	96	2	US-09-645-593-9
9	172.5	6.2	626	2	US-09-106-872A-4
10	167	6.0	524	2	US-09-424-283-1
11	164	5.9	454	2	US-09-805-694B-4
12	160	5.8	448	2	US-09-323-195A-18

ALIGNMENTS

RESULT 1

US-09-805-694B-16
; Sequence 16, Application US/09805694B
; Patent No. 6864362
; GENERAL INFORMATION:
; APPLICANT: Kinney, Anthony
; TITLE OF INVENTION: Hypoallergenic Transgenic Soybeans
; FILE REFERENCE: B01432 US NA
; CURRENT APPLICATION NUMBER: US/09/805,694B
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/189,823
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Glycine max
US-09-805-694B-16

Alignment Scores:
Pred. No.: 4.78e-133 Length: 495
Score: 1466.00 Matches: 288
Percent Similarity: 70.1% Conservative: 73
Best Local Similarity: 55.9% Mismatches: 104
Query Match: 52.8% Indels: 50
DB: 2 Gaps: 6

US-10-728-323-3 (1-1524) x US-09-805-694B-16 (1-495)

QY 1 CGCGAGCAACCGGAGGAGAACCGGTGCCAGTTCACATCGCGAGAGACTGAC 60
Db 23 ArGluGlnProGlnGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProGly 42
QY 61 AATCGCATTAATCAGAGCGGTTTACATTGAGATTGGAACCCCAACACCGAGGATTC 120

Sequence 17, Appl
Sequence 23, Appl
Sequence 17, Appl
Sequence 2, Appl
Sequence 22, Appl
Sequence 81, Appl
Sequence 14, Appl
Sequence 12, Appl
Sequence 1, Appl
Sequence 89, Appl
Sequence 24, Appl
Sequence 27, Appl
Sequence 2, Appl
Sequence 95, Appl
Sequence 97, Appl
Sequence 91, Appl
Sequence 876, Appl
Sequence 21, Appl
Sequence 93, Appl
Sequence 94, Appl
Sequence 94, Appl
Sequence 1280, Appl
Sequence 25544, Appl
Sequence 4, Appl
Sequence 17707, Appl
Sequence 20186, Appl
Sequence 4, Appl
Sequence 3, Appl
Sequence 31974, Appl
Sequence 9, Appl
Sequence 364, Appl
Sequence 7819, Appl
Sequence 2, Appl

```
Db 43 AsnArgIleGluSerGluGlyGlyLeuIleGluThrTrpAsnProAsnAsnLysProPhe 62
Qy 121 GAATGCGCGCGTGCCTCTCTCGCTTAGTCTCTCGCGCGCAACGCCCTTCGTAGCCCT 180
Db 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 82
Qy 181 TTCTACTCCAACTCTCCCAAGGAGATCTTCATCCAGCAGGAGGGGATACTTTGGTTG 240
Db 83 SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyLysGlyIlePheGlyMet 102
Qy 241 ATATTCCTCGTTGTCCTAGACACTATGAAGAGCCTCACACAAAGTCGTCCATCTCAG 300
Db 103 IleTyrProGlyCysSerThrPheGluGluProGlnGlnProGlnGlnArgGlyGln 122
Qy 301 TCCCAAAGACCAACAAGACCTCTCCAAGGAGAGACCAAGCCACAGCAACGAGATAGT 360
Db 123 SerSerArgPro-----GlnAspArg 129
Qy 361 CACAGAGGTGCACCGTTTCGATGAGGTGATCTCATTTGCAGTCCCAACCGGTGTGCT 420
Db 130 HisGlnLysIleTyrAsnSerArgGluGlyAspLeuIleAlaValProThrGlyValAla 149
Qy 421 TTCTGGCTCTACAAACGACACGACACTGATGTTGTTGCTGTTCTCTTACTGACACCAAC 480
Db 150 TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleIleAspThrAsn 169
Qy 481 AACAAACCAACAGCTGTGATCAAGTCCCGAGGATTCATTTGGCTGGGAACAACGGAG 540
Db 170 SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 189
Qy 541 CAAGAGTCTTCAAGTACCAAGCAACAAGCAGACAAAGCAGCAAGAGCAGCAAGACCTTACCATAT 600
Db 190 GlnGluPheLeuLysTyrGlnGlnGln----- 199
Qy 601 AGCCCATACAGCCGCAAGAGTCAAGCTAGACAAAGAGCGTGAAATTTAGCCCTCGAGGA 660
Db 200 -----GlyGly 201
Qy 661 CAGCACAGCGCAGAGAACGAGCAGGACAAGAGAAGAAACGAAGTGGAAACATCTTC 720
Db 202 HisGlnSerGlnLysGlyLysHisGlnGlnGluGluAsnGluGlyGlySerIleLeu 221
Qy 721 AGCGCTTACCGCGGATTCCTGGACAAAGCCTTCCAGGTTCCAGCACACACAGATAGTG 780
Db 222 SerGlyPheThrLeuGluPheLeuGluHisAlaPheSerVal---AspLysGlnIleAla 240
Qy 781 CAAAACTTAAGAGCGCAGACGAGAGTGAAGAGAGGAGGCCATTTGTGACAGTGAAGGGA 840
Db 241 LysAsnLeuGlnGlyGluAsnGluGlyGluAspLysGlyAlaIleValThrValLysGly 260
Qy 841 GGCTCTAGAATCTTGAACCCA-----GATAGAAAGAGACGTGCCGACGAGAAAGAG 891
Db 261 GlyLeuSerValIleLysProProThrAspGluGlnGlnArgProGlnGluGluGlu 280
Qy 892 GAATACCATGAAGATGATATGAATAC-----GATGAAGAG---GATAGAAGG 936
Db 281 GluGluGluGluAspGluLysProGlnCysLysGlyLysAspLysHisCysGlnArgPro 300
Qy 937 CGTGGCAGGGAAGCAGAGCAGGCGGGAATGGTATTGAAGAGACGATCTGCACCGCAAGT 996
Db 301 ArgGlySerGlnSerLysSerArgArgAsnGlyIleAspGluThrIleCysThrMetArg 320
Qy 997 GCTAAAGAAGACATTTGTAGAAACAGATCCCTGACATCTACAAACCTCAAGCTGGTTCA 1056
Db 321 LeuArgHisAsnIleGlyGlnThrSerSerProAspIleTyrAsnProGlnAlaGlySer 340
Qy 1057 CTCAAAATGCAACGATCTCAACCTCTTAATCTTAGTGGGCTTGGACCTAGTGCTGAA 1116
Db 341 ValThrThrAlaThrSerLeuAspPheProAlaLeuSerTrpLeuArgLeuSerAlaGly 360
Qy 1117 TATGGAATCTCTACAGGAATGATTTGTTGCTCGCTCACTACACCAACGACACAGC 1176
Db 361 PheGlySerLeuArgLysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSer 380
```

```
Qy 1177 ATCATATATCGATTGAGGGACGGGCTCAGTCCAGTCTGTCAGTCCGACAGCAACGCAACAGA 1236
Db 381 IleIleTyrAlaLeuAsnGlyArgAlaLeuIleGlnValValAsnCysAsnGlyGluArg 400
Qy 1237 GTGTACACGAGGAGCTTCAAGAGGTCACGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1296
Db 401 ValPheAspGlyGluLeuGlnGluGlyArgValLeuIleValProGlnAsnPheValVal 420
Qy 1297 GCTGGAAGTCCCGAGCAGAGCAATTCGAATACGTGGCATTCGAAGACAGACTCAAGGCC 1356
Db 421 AlaAlaArgSerGlnSerAspAsnPheGluTyrValSerPheLysThrAsnAspThrPro 440
Qy 1357 AGCATAGCCAACTCCCGGTGAAACTCCGTCATAGATAAATCTCCGCGGAGGAGGTGTT 1416
Db 441 MetIleGlyThrLeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIle 460
Qy 1417 GCAAAATTCATATGCGCTCCAAAGGAGCAGGCAAGCAGCTTAAGAACAAACACCCCTTC 1476
Db 461 GlnHisThrPheAsnLeuLysSerGlnGlnAlaArgGlnIleLysAsnAsnAsnProPhe 480
Qy 1477 AAGTTCTTCTCCACCGTCTCAGCAGTCTCCGAGGGCTGTGGCT 1521
Db 481 LysPheLeuValProGlnGlnGluSerGlnLysArgAlaValAla 495
```

RESULT 2

```
US-09-462-720-2
; Sequence 2, Application US/09462720
; Patent No. 6617433
; GENERAL INFORMATION:
; APPLICANT: SOCIETE DES PRODUITS NESTLE
; APPLICANT: Marraccini, Pierre
; APPLICANT: Rogers, John
; TITLE OF INVENTION: COFFEE STORAGE PROTEINS
; FILE REFERENCE: 8265-303
; CURRENT APPLICATION NUMBER: US/09/462,720
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: PCT/EP98/04038
; PRIOR FILING DATE: 1997-07-12
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Plant
US-09-462-720-2
```

```
Alignment Scores:
Pred. No.: 1,3e-69 Length: 492
Score: 811.00 Matches: 182
Percent Similarity: 52.4% Conservative: 89
Best Local Similarity: 35.2% Mismatches: 172
Query Match: 29.2% Indels: 74
DB: Gaps: 12
```

US-10-728-323-3 (1-1524) x US-09-462-720-2 (1-492)

```
Qy 1 CGCAGCAACCCGAG-----GAGAACCGTGCAGTTCAGCGCCTCAATGGCAG 51
Db 25 ArgProGlnProArgLeuArgGlyLysThrGlnCysAspIleGlnLysLeuAsnAlaGln 44
Qy 52 AGACCTTCACAATCGCATTGAATCAGAGGGGGTTACATGAGACTTGGAAACCCCAACAAC 111
Db 45 GluProSerPheArgPheProSerGluAlaGlyLeuThrGluPheTrpAspSerAsn 64
Qy 112 CAGAGTTCGAATGCGCGCGCTCGCCCTCTCTCGTCTAGTCTCCGCCCAACGCCCTT 171
Db 65 ProGluPheGlyCysAlaGlyValGluPheGluArgAsnThrValGlnProLysGlyLeu 84
Qy 172 CGTAGGCTTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAGNAGGAGGGATAC 231
Db 85 ArgLeuProHisTyrSerAsnValProLysPheValTyrValGluGlyThrGlyVal 104
```

```
QY 232 TTGGTGTGATATTCCTCGTGTCTCTAGACACTATGAAGAGCCTCACACAAAGGTCTGT 291
Db   |||  ::  |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
105 GlnGlyThrValIleProGlyCysAlaGluThrPheGlu-----SerGlnGlyGlu 121
QY 292 CGA-----TCTCAGTCCCAAGACACCAAGACGTCTCAAGGAGAACACAA----- 339
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
122 SerPheTrpGlyGlnGluProGlyLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 141
QY 340 -----AGCCAAACAGCAACGA-----GATAGTCACCAGAAAGGTGCACCGT 378
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
142 LysGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 161
QY 379 TTCGATGAGGGTATCTCATGCGATTCCCAACCGGTGTTGCTTCTGCTCTACACGAC 438
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
162 PheGlnLysGlyAspValLeuLeuLeuLeuProGlyPheThrGlnTrpThrTrpAsnAsp 181
QY 439 CAGCAGCTGATGTTGTTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 498
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
182 GlyAspValProLeuValThrValAlaLeuLeuAspValAlaAsnGluAlaAsnGlnLeu 201
QY 499 GATCAGTTCCTCCAGAGATTCAATTTGGCTGGGAACACGAGCAAGAGTCTTAAGGTAC 558
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
202 AspLeuGlnSerLysLysPheLeuAlaGlyAsnProGlnGlnGlnGlyGlyLysGlu 221
QY 559 CAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATATAGCCCATACAGCCCGCAA 618
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
222 GlyHisGlnGlnGlnGlnHisArg----- 231
QY 619 AGTCAGCTAGACAAGAGAGCGTGAATTTAGCCCTCGAGGACGACGACCGCGCAGAA 678
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
231 ----- 231
QY 679 CGAGCAGCACAAGAAGAAAGAAAGGTTGGAACATCTTCAGCGGCTTCACGCGGAG 738
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
232 -----AsnIlePheSerGlyPheAspGln 240
QY 739 TTCTCTGAACAAGCTTCAGGTTGACGACAGACAGATAGTCAAAACCTAAGAGCGGAG 798
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
241 LeuLeuAlaAspAlaPheAsnVal---AspLeuLysIleIleGlnLysLeuLysGly--- 258
QY 799 ACCGAGAGTGAAGAGAGGAGCCATTGTGACAGTGAAGGGA---GGCTCAGATCTTG 855
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
259 ---ProLysAspGlnArgGlySer-----ThrValArgAlaGluLysLeuGlnLeuPhe 275
QY 856 AGCCACAGATAGAAGAGAGCTGCCGACGAAAGAGAGAAATACGATGAATGAATGAA 915
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
276 LeuProGlyLysSerGluGlnValGlnGlnProGlnGln----- 288
QY 916 TAGCATGAAGAGATAGAAGCGGTGCGAGGGAAGCAGAGGCGGGGAATCGTATTGAA 975
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
289 ---GlnGlnGlnGlnGlnHisGlyValGlyArgGlyTrpArgSerAsnGlyLeuGlu 307
QY 976 GAGACGATCTGACCGCAAGTCTAAAGAAACATTTGTTAGAAACAGATCCCTCAGATC 1035
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
308 GluThrLeuCysThrValLysLeuSerGluAsnIleGlyLeuProGlnGlnAlaAspVal 327
QY 1036 TACAACCTCAAGCTGTTCACTCAAACTGCAACAGCTCTCAACCTCTTAATCTTAGG 1095
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
328 PheAsnProArgAlaGlyArgIleThrThrValAsnSerGlnLysIleProIleLeuSer 347
QY 1096 TGGCTTGGACCTAGTCTGAAATATGGAATCTCTACAGGAATGCAATGTTTCTGCTCAC 1155
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
348 SerLeuGlnLeuSerAlaGluArgGlyPheLeuTrpSerAsnAlaIlePheAlaProHis 367
QY 1156 TACAACACCAACGACACAGCATCATATATTCATTGAGGGGACGGCTCACGTGCAATGC 1215
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
368 TrpAsnIleAsnAlaHisAsnAlaLeuTrpValIleArgGlyAsnAlaArgIleGlnVal 387
QY 1216 GTGGCAGCAACGCGCAAGAGTGTACGACGAGGAGCTTCAAGAGGTCACGTGTTGTG 1275
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
388 ValAspHisLysGlyAsnLysValPheAspAspGluValLysGlnGlnGlyGlnLeuIle 407
QY 1276 GTGCCACAGAACTTCGCGCGTCTGGAAAGTCCCAGAGCGAGAACTTCGAATACGTGGCA 1335
```

```
Db 408 ValProGlnTrpPheAlaValIleLysLysAlaGlyAsnGlnGlnGlyPheGluTrpValAla 427
QY 1336 TTCAGACAGACTCAAGCCAGCATACCCACCTCGCGGTGAAAACCTCCGTCTAGAT 1395
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
428 PheLysThrAsnAspAsnAlaMetIleAsnProLeuValGlyArgLeuSerAlaPheArg 447
QY 1396 AACCTCGCGGAGGAGTGTTCGCAATTCATATGCGCTCCAAAGGGAGCAGGCAAGGAC 1455
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
448 AlaIleProGluGluValLeuArgSerPheGlnIleSerSerGluGluAlaGluGlu 467
QY 1456 CTTAAGAACAACAACCCCTTCAAGTTCCTGTTCCACCGCTCTCAGCAGTCT 1506
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
468 LeuLysTrpGlyArgGlnGluArgLeuLeu-----SerGluGlnSer 482

RESULT 3
US-10-053-410-6
; Sequence 6, Application US/10053410
; Patent No. 6858778
; GENERAL INFORMATION:
; APPLICANT: Sewalt, Vincent J. H.
; APPLICANT: Jung, Rudolf
; APPLICANT: Hu, Wang-Nan
; APPLICANT: Meeley, Robert B.
; TITLE OF INVENTION: Improved Grain Quality Through Altered
; FILE REFERENCE: Expression of Seed Proteins
; CURRENT APPLICATION NUMBER: US/10/053,410
; CURRENT FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,455
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Zea mays
US-10-053-410-6

Alignment Scores:
Pred. No.: 1,86e-51 Length: 483
Score: 623.50 Matches: 151
Percent Similarity: 45.1% Conservative: 77
Best Local Similarity: 29.9% Mismatches: 172
Query Match: 22.4% Indels: 105
DB: 2 Gaps: 11

US-10-728-323-3 (1-1524) x US-10-053-410-6 (1-483)
QY 25 TGCCAGTTTCCAG---CGCTCAATGCGCAGAGACCTGACAATCGCATTTGAATCAGAGGCG 81
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
48 CysGlyPheAspGlyLysLeuGluAlaLeuGluProArgHisLeValGlnSerGluAla 67
QY 82 GGT-----TACATTGAGACTTGGAAACCCCAACACACAGGAGTTTGAATGCCCGCGC 132
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
68 GlySerValGlnTrpPheSerArgPheAsnGluAlaAspArgGluLeuThrCysAlaGly 87
QY 133 GTCGCCCTCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGSCCTTTCTACTCCAAT 192
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
88 IlePheAlaValArgValValValAspAlaMetGlyLeuLeuLeuProArgTrpSerAsn 107
QY 193 GCTCCCGCAGGAGATCTTCATCCAGCAAGGAGGGGATACTTTGGGTGTGATATTCCTCGT 252
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
108 ValHisSerLeuValTrpIleValGlnGlyArgGlyIleIleGlyPheSerPheProGly 127
QY 253 TGCTCTAGACACTATGAAGAGCCTCACACAAAGTCGTCGATCTCAGTCCCAAGACCA 312
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
128 Cys----- 128
QY 313 CCAAGACGCTCTCCAAGAGAGAACCAAGCCCAACAG----- 348
Db   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
129 -----GlnGluGluThrGlnGlnGlnTrpGlyGlyGlyGlyGlyGlyGly 144
```


; APPLICANT: Moloney, Maurice
; APPLICANT: Singh, Surinder
; TITLE OF INVENTION: Flax Seed Specific Promoters
; FILE REFERENCE: 9369-151
; CURRENT APPLICATION NUMBER: US/09/645,593
; CURRENT FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/151,044
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: US 60/161,722
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Linum usitatissimum
; SEQ ID NO 10

Alignment Scores:
Pred. No.: 1,69e-13 Length: 85
Score: 227.50 Matches: 46
Percent Similarity: 61.2% Conservative: 14
Best Local Similarity: 46.9% Mismatches: 25
Query Match: 8.2% Indels: 13
DB: 2 Gaps: 2

US-10-728-323-3 (1-1524) x US-09-645-593-10 (1-85)
QY 220 GGAAGGGATACCTTGGGTGTGATATTCCTCGTGTGTCTAGACACTATGAAGAGCTCTAC 279
Db 1 GlyArgGlyValThrGlyIleMetPhePro**CysProGluThrPheGluGluSerGln 20
QY 280 ACACAGGTCGTCGATCTCAGTCCCAAGACCACCAAGACGCTCCACGAGAACACCAA 339
Db 21 GlnGlnGlyGlnGlnGlnGln-----GlnGlySerSerGln 32
QY 340 AGCCACAGACGAGATAGTCACAGAAAGGTGCACCGTTTCGATGAGGGTGATCTCATT 399
Db 33 -----AspGlnHisGlnLysIleArgArgPheArgGluGlyaspValIle 47
QY 400 GCAGTTCACCGGTGTTCTTCTGCTCTACACGACACACGACACTGATGTTGTGCT 459
Db 48 AlaValProAlaGlyValAlaHisTrpSerTyrAsnAspGlyAsnGluProValMetAla 67
QY 460 GTTCTCTTACTGACACCAACACACGACACCGCTTGATCAGTTCCTCCAGG 513
Db 68 IleValValHisAspThrSerSerHisLeuAsnGlnLeuAspAsnProArg 85

RESULT 6
US-09-645-593-11
; Sequence 11, Application US/09645593
; Patent No. 6777591
; GENERAL INFORMATION:
; APPLICANT: Chaudhary, Sarita
; APPLICANT: van Rooijen, Gijb
; APPLICANT: Moloney, Maurice
; APPLICANT: Singh, Surinder
; TITLE OF INVENTION: Flax Seed Specific Promoters
; FILE REFERENCE: 9369-151
; CURRENT APPLICATION NUMBER: US/09/645,593

; CURRENT FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/151,044
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: US 60/161,722
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Linum usitatissimum
; SEQ ID NO 11

Alignment Scores:
Pred. No.: 3.89e-12 Length: 165
Score: 215.00 Matches: 62
Percent Similarity: 43.1% Conservative: 28
Best Local Similarity: 29.7% Mismatches: 69
Query Match: 7.7% Indels: 50
DB: 2 Gaps: 6

US-10-728-323-3 (1-1524) x US-09-645-593-11 (1-165)

QY 517 TTCATTTGGCTGGGAACACGAGCAAGAGTTCTTAAGGTACCAACAAAGCAGACAA 576
Db 2 PheTyrLeuAlaGlyAsnProArgAspGluPheGluGlnSerGlnGlnGlyGlyArgLeu 21
QY 577 AGCAGACCAAGAGCTTACCATATAGCCATACAGCCCGCAAGTCAAGCTAGACAA 636
Db 22 SerArgGlyGluSer----- 26
QY 637 GAGCGTGAATTTAGCCCTCGAGGACAGCACAGCCGACGAGAGCAGGACCAAGAAGAA 696
Db 27 -----GluGlyGlyArgGlyArgGluProLeuGlnProAlaThr 40
QY 697 GAAACAGAGGTGGAAACATTTACGCGCTTACGCGGAGTTCTCGAACCAAGCCTTC 756
Db 41 ThrSerSer-----CysGlyIleAspSerLysLeuIleAlaGluAlaPhe 55
QY 757 CAGGTTGACGACAGACATAGTCAAAACCTAAGAGCGGAGACGAGAGTGAAGAGAG 816
Db 56 AsnValAspGlu---AsnValAlaArgArgLeu-----GlnSerGluAsnAspAsnArg 72
QY 817 GGAGCCATTGTGACAGTGAAGGAGGCGCTCAGAATCTTGAGCCGACAGATAGAAAGAGACGT 876
Db 73 GlyGlnIleValArgValGluGlyGluLeuAspIleValArgProThrSerIleGln 92
QY 877 GCCGACGAGAGAGGAATACGATGAAGATGAATATGAATACGATGAAGAGGATGAAGG 936
Db 93 GluGluSerGlnGluGln----- 98
QY 937 CGTGCAGGGAAGCAGAGGAGGAGGAGG-----AATGGTATTGAGAGAGCAGATC 984
Db 99 -----GlyGlyArgGlyGlyGlyArgTyrTyrSerAsnGlyValGluGluThrPhe 115
QY 985 TGCACCGCAAGTGTCTAAAAGAAACATTCGTAGAAACAGATCCCTCGACATCTTACACCT 1044
Db 116 CysSerMetArgLeuIleGluAsnIleGlyAspProSerArgAlaAspIlePheThrPro 135
QY 1045 CAAGCTGTTCTCACTCAAAACCTGCCAACGATCTCAACCTTCTTAATCTTAGTGGCTTGA 1104
Db 136 GluAlaGlyArgValArgSerLeuAsnSerHisAsnLeuProValLeuGlnTrpIleGln 155
QY 1105 CTAAGTCTGAATATGGAATCTCTAC 1131
Db 156 LeuSerAlaGluArgGlyValLeuTyr 164

RESULT 7
US-09-424-283-3
; Sequence 3, Application US/09424283
; Patent No. 6437219
; GENERAL INFORMATION:
; APPLICANT: Grimes, et al.

QY	682	GCAGGCAAGAAGAAACAGAGGTGGTGAACATCTTCAGCGCTTCACGCCGAGTTC	741
Db	233	::: ::::: ::: --AspProGluSerValIleuSerAlaPheSerTrpAsnVal	245
QY	742	CTGGAAACAAAGCCTTCAGGTTGACGACGACAGATAGTGTCAAAACCTTAAGAGCGGACAC	801
Db	246	: -----LysGlyLysLeu	257
QY	802	GAG-----AGTGAAGAAGAGGAGCCATTGTGACAGTG-----AGGGAGGCGCTC	846
Db	258	GluArgLeuPheAsnGlnGlnAsnGluGlySerIlePheLysIleSerArgGluArgVal	277
QY	847	AGAACTCTGAGCCCATAGAAAGACAGCGTGCACGACGAAGAAGAGGAATACGATGAAGAT	906
Db	278	ArgAlaLeuAlaProThrLysLys-----SerSer	287
QY	907	GRATTAAGTAATCGATGAAGAGGATAGAAGCGTGGCAGGGGAACGAGAGCGAGGGGAAT	966
Db	288	TrpTipProPheGlyGlyGluSerLys-----	296
QY	967	GGTATTGAAGACAGCATCTCCACCGCAGTGCTAAAAAGAACATTGGTAGAGAAACAGATCC	1026
Db	297	-----AlaGlnPheAsnIlePheSerLysArg-----	305
QY	1027	CCTGACATCTACAACCCCTCAAGCTGGTTCACTCAAAACCTCCCAACGATCTCAACCTTCTA	1086
Db	306	ProThrPheSerAsnGlyTyArgLeuThrGluValGlyProAspGluLysSer	325
QY	1087	ATACTTAGTGGCTTGACCTAGTGTGAATATGGAAATCTCTACAGGAATGCATTGTTT	1146
Db	326	TrpLeuGlnArgLeuAsnLeuMetLeuThrPheThrAsnIleThrGlnArgSerMetSer	345
QY	1147	GTCGCTCACTACAACACCAACGACACAGCATCATATATCGATTGAGGGCAGCGGCTCAC	1206
Db	346	ThrIleHisTyAsnSerHisAlaThrLysIleAlaLeuValMetAspGlyArgGlyHis	365
QY	1207	GTGCAAGTC-----GTGGACAGCAACGGCAACAGAGTGTCACGACGAG-----	1248
Db	366	LeuGlnIleSerCysProHisMetSerArgSerAspSerLysHisAspLysSerSer	385
QY	1249	-----GAGCTTCAAGAGGGTCACTGCTGCTGTTGGTGGTCCACAG	1284
Db	386	ProSerTyrHisArgIleSerAlaAspLeuLysProGlyMetValPheValValProPro	405
QY	1285	AACTTCGCC---GTCGCTGGAAGATCCAGAGCGCAGAACTTCGAATACGTGGCATTCAG	1341
Db	406	GlyHisProPheValThrIleAlaSerAsnLysGluAsnLeuLeuIleCysPheGlu	425
QY	1342	ACAGACTCAAGGCCACCATAGCC---AACTGCGCGGTGAAAACCTCCGTCATA-----	1399
Db	426	ValAsnValArgAspAsnLysLysPheThrPheAlaGlyLysAspAsnIleValSerSer	445
QY	1393	---GATAAATCTGCGGAGGAGGTGGTTGCAAAATTTCATATGCGCTCCAAAAGGAGCAGGCA	1449
Db	446	LeuAspAsnValAlaLysGluLeuAlaPheAsnTyrProSerGluMetValAsnGlyVal	465
QY	1450	AGGCAGCTTAAGAACACAAACCCCTTCAAGTCTCTTCGCTTCACCGCTCTCAGCAGTCTCCG	1508
Db	466	SerGluArgLysGluSerLeuPhePheProPheGluLeuProSerGluGluArgGlyArg	485
QY	1510	AGGCGTGTGGCT	1521
Db	486	ArgAlaValAla	489

RESULT 8

US-09-645-593-9

; Sequence 9, Application US/09645593

; Patent No. 6777591

; GENERAL INFORMATION:

; APPLICANT: Chaudhary, Sarita

; APPLICANT: van Rooijen, Gijls

; APPLICANT: Moloney, Maurice


```
Db 164 ----- 164
QY 529 GGAACACGAGCAGAGTCTTAAAGTACCAGCAACAAAGCAGACAGCAAGA 588
Db 164 ----- 164
QY 589 AGCTTACCATATAGCCCATACAGCCGCAAGTCAAGCTAGACAAAGAGCGGTGAATTT 648
Db 165 ----- 165
QY 649 AGCCCTCGAGGACAGCAGCCGACAGACGAGCAGGACAGACAGAAACAGAGT 708
Db 181 TyrValAlaGly ----- 189
QY 709 GGAACATCTTACGCGCTTACGCGGAGTTCTCGAACACAGCTTCCAGGTTGACGAC 768
Db 190 ---ThrValTyrSerAlaPheSerAspValLeuGluAlaAlaPheAsnThrAsnVal 208
QY 769 AGACAGATAGTCGAAACCTAAGAGCGGACCGAGAGTGAAGAGAGGAGCGCATTTGTG 828
Db 209 GlnGlnLeuGluArgIlePheGlyGly ----- HisLysSerGlyValIleIle 224
QY 829 ACAGTCAGGGGCGCTCAGATCTTGAGCCGACAGATCTGACCGCAGTCTAAAGAGAAC 888
Db 225 ----- HisAlaAsnGluGlu 229
QY 889 GAGGAATACGATGAAGTGAATATGATGATGATGATGATGATGATGATGATGATGAT 948
Db 230 ----- GlnIleArgGluMetCysArgLysArgGly ---Phe 240
QY 949 AGCAGAGCGAGGGGAGTATTTGAAGAGACGATCTGACCGCAGTCTAAAGAGAAC 1008
Db 241 SerAlaGlySerMetSerAlaProGluHis ----- ProLysProPhe 254
QY 1009 ATTGGTAGAACAGATCCCTGACATCTACACCTCAAGCTGCTCACTCAAAAGTGC 1068
Db 255 AsnLeuArgAsnGlnLysProAspPheGluAsn ---GluAsnGlyArgPheThrIleAla 273
QY 1069 AACGATCTCAACCTTCTAATATCTTAGTGTGGCTAGTGTCTGAATATGAAATCTC 1128
Db 274 GlyProLysAsnTyrProPheLeuAspAlaLeuAspValSerValGlyLeuAlaAspLeu 293
QY 1129 TACAGGAATGCATTTGTCGCTCACTACACCAACGACACGACGATCATATATCGA 1188
Db 294 AsnProGlySerMetThrAlaProSerLeuAsnSerLysSerThrSerIleGlyIleVal 313
QY 1189 TTGAGGGGAGGGCTCAGCTGCAAGTCTGTG ----- 1218
Db 314 ThrAsnGlyGluGlyArgIleGluMetAlaCysProHisLeuGlyGlnHisGlyTrpSer 333
QY 1219 -----GACAGCAACGGCAACAGAGTGTACGACGAGGCTT 1254
Db 334 SerProArgGluArgGlyAspGlnAspIleThrTyrGlnArgValTrp ---AlaLysLeu 352
QY 1255 CAAGAGGGTCAGCTGCTGTGGTGCACAGAACTTCGCGCTCGCTGGA ---AAGTCCACAG 1311
Db 353 ArgThrGlySerValTyrIleValProAlaGlyHisProIleThrGluIleAlaSerThr 372
QY 1312 AGCGAGAATCTCGAATACCTGGCATTCAGACAGACTCAAGCCCGACGATACCCAAC --- 1368
Db 373 AsnSerArgLeuGlnIleLeuTrpPheAspLeuAsnThrArgGlyAsnGluArgGlnPhe 392
QY 1369 CTCGCGGTGAAACCTCCGTCATAGATAACCTTCGCGGAGGAGTGTGTTCAAAATTCATAT 1428
Db 393 LeuAlaGlyLysAsnAsnValLeuAsnThrLeuGluArgGluIle ----- 407
QY 1429 GGCCTCCAAAGGAGCAGCAAGGCGAGCTTAAGAACAAACACCC 1473
Db 408 -----ArgGlnLeuSerPheAsnValPro 415
```

RESULT 13

US-09-323-195A-17

```
; Sequence 17, Application US/09323195A
; Patent No. 6462257
; GENERAL INFORMATION:
; APPLICANT: Pullman, Gerald
; APPLICANT: Cairney, John
; APPLICANT: Ferreira, Ranjan
; TITLE OF INVENTION: VICILIN-LIKE SEED STORAGE PROTEIN GENE PROMOTER AND
; FILE REFERENCE: IPST0009
; CURRENT APPLICATION NUMBER: US/09/323,195A
; CURRENT FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Pinus taeda
US-09-323-195A-17
```

```
Alignment Scores:
Pred. No.: 2,35e-06 Length: 523
Score: 158.00 Matches: 107
Percent Similarity: 32.7% Conservative: 74
Best Local Similarity: 19.3% Mismatches: 190
Query Match: 5.7% Indels: 183
DB: 2 Gaps: 20
```

US-10-728-323-3 (1-1524) x US-09-323-195A-17 (1-523)

```
QY 10 CCGAGGAGAGAACCGCTGCCAGTTCAGCGCCTCAATGCGCAGACCTGACAAAT----- 63
Db 37 ProGluAspHisGlyArgGlyHisGlnArgArgGluGluGluArgGluGluAsnProTyr 56
QY 64 -----CGCATTGAATCAGAGGGCGGTTACATTGAGACT 96
Db 57 ValPheHisSerAspArgPheArgMetArgAlaSerSerAspAlaGlyLeuIleArgAla 76
QY 97 TGGAAACCCCAACCAACAGGAGTTCGAA-----TGC CGCGCGCTCGCCCTCTCTCGCTTA 150
Db 77 Leu---ProAsnPheGlyGluAlaSerGluLeuLeuGluGlyIleSerLysTyrArgVal 95
QY 151 GTC-----CTCGCGCGCAACGCCCTTCGTAGCGCTTCTACTCCAAATGCTCCCCAG 201
Db 96 ThrCysIleGluMetArgProAsnThrValMetLeuProHisTyrLeuAspAlaThrTrp 115
QY 202 GAGATCTTCATCCAGCAAGGAGGATACTTTGGGTTGATATTCCTGGTTGTCCTAGA 261
Db 116 IleLeuTyrValThrGlyArgGlyTyrIleAlaTyrVal----- 129
QY 262 CACTATGAAGAGCCTCACACACAGGTCGTCTCATCTCAGTCCCAAGACCAACAGACGT 321
Db 130 -----HisGlnAsnGluLeuValLysArgLys 138
QY 322 CTCCAAGGAGAGAACCAAGCCACGACGACGACGACGACGACGACGACGACGACGACG 381
Db 139 Leu----- 139
QY 382 GATGAGGTGATCTCATTCGAGTTCACCGGTGTTGCTTCTGCTCTACAACGACAC 441
Db 140 GluGluGlyAspValPheGlyValProSerGlyHisThrPheTyrLeuValAsnAsnAsp 159
QY 442 GACACTGATGTTGTTGCTGTTTCTCTTACTTGACACCAACCAACGACCAACGACCTTGAT 501
Db 160 AspHisAsnSerLeuArgIleThr----- 167
QY 502 CAGTTCCTCCAGGAGATTCATTTGGCTGGGAACACGGAGCAAGATTCCTTAAGGTAC 561
Db 168 -----SerLeuLeuArgThrVal 173
QY 562 CAACAAAGCAGACAAAGCAGACGACGACGACGACGACGACGACGACGACGACGACG 621
Db 174 SerThrMetArgGlyGlu-----TyrGluProTyr----- 183
```

```
QY 622 CAGCCTAGACAAGAGCGTGAATTTAGCCCTCGAGGACAGCAGCAGCCGCGAGAACGA 681
Db 184 -----TyrValAla 186
QY 682 GCAGGACAAGAAGAAACGAAGGTGGAACATCTTCAGGGCTTTCAGCGCGGAGTTC 741
Db 187 GlyIArgAsnProGlu-----ThrValIySerAlaPheSerIaspVal 202
QY 742 CTGGAACAAGCCTTCAGGTTGACGACAGACAGATAGTGCAAAACCTAAGAGCGAGACC 801
Db 203 LeuGluAlaAlaPheAsnThrAsn----- 210
QY 802 GAGAGTGAAGAGAGGAGCCATTGTGACAGTCAGGGGAGGCTCAGATCTTTGACCCCA 861
Db 211 -----ValIleGluAlaArgThrHisPheProVal 220
QY 862 GATAGAAAGAGCGTCCGACGAAGAAGAGGAATACGATGACAGATGAATATGATACGAT 921
Db 221 -----HisIleGluArgGluSerTyrSerMetAla 230
QY 922 GAAGAGATAGAAAGCGTGGCAGGGGAAGCAGAGGCGAGGGGAATGTTATTGAAGAGCG 981
Db 231 AsnGluGluGlnIleArg--GluMetLeuArgLysArgGlyPheSerAlaGluSerMet 249
QY 982 ATCTGCACCGCAAGTCTCTAAAGAACAATGTTAGAAACAGATCCCTGCATCTACAAAC 1041
Db 250 SerAlaSerGluHisProLysProPheAsnLeuArgAsnGlnLysProAspPheGluAsn 269
QY 1042 CCTCAAGCTGGTTCACCTCAAAACCTCAACCTCTCAATCTTAACTAGTGGCTT 1101
Db 270 AspAsnGlyArgPheThrArgAlaGlyProAsnGluAsnProLeu---LeuAspAlaVal 288
QY 1102 GGACCTAGTCTGAATATGAAATCTCTACAGGAATGCAATGTTGTCTGCTCACTACAAC 1161
Db 289 AspValThrAlaGlyPheGlyValLeuAsnProGlyThrMetThrAlaProSerHisAsn 308
QY 1162 ACCAAGCCACAGCATCATATCATGATTGAGGGGAGCGGCTCACGTGCAAGTCGTG--- 1218
Db 309 ThrLysAlaThrSerIleAlaIleValThrGlnGlyGluArgIleGluMetAlaCys 328
QY 1219 -----GACAGCAAC 1227
Db 329 ProHisLeuGlyGlnHisGlyTyrSerArgArgGluLysGlyAspGlnGluIleAsn 348
QY 1228 GGCAACAGAGTGTPACGACGAGGAGCTTCAAGAGGTGTCAGTGTGTGTGTCACAGAAC 1287
Db 349 TyrGlnArgVal---ArgAlaArgLeuArgThrGlyThrValTyrValProAlaGly 367
QY 1288 TTCGCGCTCGCTGGAAGTCCAGAGCGAG--AACTTCGAATACGTGGCATTCAGACA 1344
Db 368 HisProIleThrGluIleAlaCysThrGluGlyHisLeuGluIleLeuTrpPheAspIle 387
QY 1345 GACTCAAGGCCACGATAGCCAAC---CTGCGCGGTGAAACCTCCGTCATAGATAACCTG 1401
Db 388 AsnThrSerGlyAsnGluArgGlnPheLeuAlaGlyLysTyrAsnValLeuGlnThrLeu 407
QY 1402 CCGAGAGAGTGGTTCGAAATTCATATGCGCTCCAAAGG-----GAGCAG 1446
Db 408 GluLysGluValArgGlnIleSerPheAsnIleProArgGlyGluGluLeuAspGluVal 427
QY 1447 GCAAGCAGCTTAAGAAC-----AACAAAC 1470
Db 428 LeuArgArgGlnLysAspGlnValIleLeuArgGlyProGlnMetGlnArgAspGlu 447
QY 1471 CCCTTCAAGTTCTTCGTTCCACCGTCTCAGCAGTCTCCGA 1510
Db 448 -ProArgSerSerIleHisAlaIleIleAlaAlaArg 460
```

RESULT 14
US-09-106-872A-17
; Sequence 17, Application US/09106872A
; Patent No. 6486311
; GENERAL INFORMATION:

```
; APPLICANT: Burks Jr., A. Wesley
; APPLICANT: Stanley, J. Steven
; APPLICANT: Cockrell, Gael
; APPLICANT: King, Nina E.
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Helm, Ricki M.
; APPLICANT: Bannan, Gary A.
; TITLE OF INVENTION: Peanut Allergens and Methods
; FILE REFERENCE: HS 103 CIP
; CURRENT APPLICATION NUMBER: US/09/106,872A
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: PCT/US96/15222
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-09-106-872A-17

Alignment Scores:
Pred. No.: 2,08e-06 Length: 335
Score: 157.50 Matches: 55
Percent Similarity: 42.0% Conservative: 48
Best Local Similarity: 22.4% Mismatches: 81
Query Match: 5.7% Indels: 61
DB: 2 Gaps: 9

US-10-728-323-3 (1-1524) x US-09-106-872A-17 (1-335)
QY 31 TTTCAGCGCTCAATCGCGCAGAGACTGACATCGCATCGAATCAGAGCGGTTACATT 90
Db 122 PheGlyLysLeuPheGluValLysProAspLysLys----- 133
QY 91 GAGACTTGGAACCCCAACCAACAGGAGTTTCGAATGCGCGGCGTCCCTCTCTCGCTTA 150
Db 134 -----AsnProGlnLeuGlnAspLeuAsp-----MetMetLeuThrCysVal 147
QY 151 GTCCTCCGCGCAACGCCCTTCGTAGGCCT---TTCTACTCCAATGCTCCCCAGGAGATC 207
Db 148 GluIleLysGluGlyAlaLeuMetLeuProHisPheAsnSerLysAlaMetValIleVal 167
QY 208 TTCATCCAGCAAGGAAGGGATACCTTGGGTTGATATTCCTGGTGTCTCTAGACACTAT 267
Db 168 ValValAsnLysGlyThrGlyAsnLeuGluLeuVal----- 179
QY 268 GAAGAGCCTCACACACAGGTGCTCGATCTCAGTCCCAAGACCAACAGACGCTCCAA 327
Db 180 -----AlaValArgLysGluGlnGlnArgGlyArgGluGlu 193
QY 328 GGAGAGACCAAGCCACAGCAACGAGATAGTACCAAGAGGTGACCGTTTC----- 381
Db 194 GluGluAspGluAspGluGluGluGlySerAsnArgGluValArgGlyThrAla 213
QY 382 -----GATGAGGTGATCTCATTGCAGTTCACCACCGGTGTGCTTCTGGCTCTACAAC 435
Db 214 ArgLeuLysGluGlyAspValPheIleMetProAlaAlaHisProValAlaIleAsnAla 233
QY 436 GACCACGACCTGATGTTGTTGCTTCTCTTCTTACTGACACCAACCAACAC----- 486
Db 234 SerSerGluLeuHisLeuLeuGlyPheGlyIleAsnAlaGluAsnHisArgIlePhe 253
QY 487 -----GACAAACAGCTTGTATGATGTTCCCGAGGAGA----- 516
Db 254 LeuAlaGlyAspLysAspAsnValIleAspGlnIleGluLysGlnAlaLysAspLeuAla 273
QY 517 TTCAATTTGGCTGGGAACACGAGGAGAGTCTTTAAGGTACCAAGCAACAAACAGACAA 576
Db 274 PheProGlySerGlyGluGlnValGluLysLeuIleLysAsnGlnLysGluSerHisPhe 293
QY 577 AGCAGACGAAGAAGCTTACCATATATAGCCCATATAGCCCGCAAGAGTGCAGCTAGACAAGAA 636
```



```
Db      428 SerIlePheValProHisTyrAsnSerLysAlaThrPheValValLeuValAsnGluGly 447
QY      1198 CGGGCTCACGTGCAAGTCGTGGAC----- 1221
      ::|||:::|
Db      448 AsnGlyTyrValGluMetValSerProHisLeuProArgGlnSerSerPheGluGluGlu 467
QY      1222 -----AGCAACGGCAACAGAGTG 1239
      ::|||
Db      468 GluGluGlnGlnGlnGluGlnGluGluGluArgSerGlyGlnTyrArgLys 487
QY      1240 TACGACGAGGAGTTCAAGAGGGTCACGTGCTTGTGTCACAGAACTTCGCCGTCGCT 1299
      ::|||
Db      488 IleArgSerGlnLeuSerArgGlyAspIlePheValValProAlaAsnPheProValThr 507
QY      1300 ---GGAAGTCCAGAGCGAGAACTTCGAATACGTGGCATTC----- 1338
      |||||:::|
Db      508 PheValAlaSerGlnAsnGlnAsnLeuArgMetThrGlyPheGlyLeuTyrAsnGlnAsn 527
QY      1339 ---AAGACAGACTCAAGGCCCATAGCCAACTCGCGGTGAAACTCCGTCATAGAT 1395
      |||
Db      528 IleAsnProAspHisAsnGlnArgIlePhe---ValAlaGlyLysIleAsnHisValArg 546
QY      1396 AACCTGCCGAGGAGGTGTTCAAATTCATATGGCTCCAAAGGGAGCAGGCAAGGCAG 1455
      ::|
Db      547 GlnTrpAspSerGlnAlaLysGluLeuAlaPheGlyValSerSerArgLeuValAspGlu 566
QY      1456 CTTAAGAACAACACCCCTTCAAG---TTCTTCGTTCCACCGTCTCAGCAGTCTCCGAGG 1512
      ::|||
Db      567 IlePheAsnAsnAsnProGlnGluSerTyrPheVal-----SerArgGlnArgGlnArg 584
QY      1513 GCT 1515
Db      585 Ala 585
```

Search completed: May 15, 2006, 22:28:33
Job time : 64.8306 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 23:34:07 ; Search time 43.013 Seconds
(without alignments)
4441.255 Million cell updates/sec

Title: US-10-728-323-3

Perfect score: 2779

Sequence: 1 cggcagcaaccggaggagaa.....ctccgagggtgtggtcttaa 1524

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_n2p.model -DEV=xlp
-Q=/abs/ABSWEB_spool/US10728323/runat_15052006_172145_22607/app_query.fasta_1
-DB=Published Applications AA Main -OFMT=fastan -SUFFIX=n2p.rapbm
-MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=US10728323@CGN 1 1_503_runa_15052006_172145_22607 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Published Applications AA Main:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp.*
 - 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pcp.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp.*
 - 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2673	96.2	507	4	US-10-245-871-96 Sequence 96, Appl
2	2673	96.2	507	4	US-10-253-286-96 Sequence 96, Appl
3	2673	96.2	510	4	US-10-228-803-6 Sequence 6, Appl
4	2673	96.2	510	4	US-10-100-306-90 Sequence 90, Appl
5	2638	94.9	526	3	US-09-731-221-79 Sequence 79, Appl
6	2621	94.3	530	5	US-10-899-551-6 Sequence 6, Appl
7	1665	59.9	351	5	US-10-899-551-58 Sequence 58, Appl
8	1502	54.0	481	3	US-09-759-967-21 Sequence 21, Appl
9	1502	54.0	481	4	US-10-424-599-171702 Sequence 171702,
10	1502	54.0	488	4	US-10-425-114-43865 Sequence 43865, A
11	1502	54.0	491	4	US-10-425-114-44047 Sequence 44047, A

12	1502	54.0	492	4	US-10-425-114-43836 Sequence 43836, A
13	1502	54.0	492	4	US-10-425-114-46294 Sequence 46294, A
14	1499.5	54.0	466	4	US-10-425-114-71959 Sequence 71959, A
15	1499.5	54.0	466	4	US-10-425-114-71973 Sequence 71973, A
16	1499.5	54.0	485	3	US-09-759-967-20 Sequence 20, Appl
17	1499.5	54.0	488	4	US-10-425-114-45645 Sequence 45645, A
18	1499.5	54.0	489	4	US-10-425-114-43983 Sequence 43983, A
19	1499.5	54.0	489	4	US-10-425-114-44676 Sequence 44676, A
20	1499.5	54.0	489	4	US-10-425-114-45612 Sequence 45612, A
21	1499.5	54.0	489	4	US-10-425-114-50117 Sequence 50117, A
22	1499.5	54.0	489	4	US-10-425-114-50207 Sequence 50207, A
23	1499.5	54.0	489	4	US-10-425-114-51412 Sequence 51412, A
24	1499.5	54.0	489	4	US-10-425-114-58429 Sequence 58429, A
25	1499.5	54.0	489	4	US-10-425-114-71902 Sequence 71902, A
26	1499.5	54.0	489	4	US-10-425-114-71909 Sequence 71909, A
27	1499.5	54.0	489	4	US-10-425-114-71957 Sequence 71957, A
28	1499.5	54.0	489	4	US-10-425-114-71958 Sequence 71958, A
29	1499.5	54.0	489	4	US-10-425-114-71967 Sequence 71967, A
30	1499.5	54.0	489	4	US-10-425-114-71995 Sequence 71995, A
31	1499.5	54.0	489	4	US-10-425-114-72287 Sequence 72287, A
32	1499.5	54.0	489	4	US-10-425-114-72288 Sequence 72288, A
33	1498	53.9	489	4	US-10-425-114-45639 Sequence 45639, A
34	1496	53.8	476	5	US-10-409-993-1 Sequence 1, Appl
35	1496	53.8	477	5	US-10-409-993-15 Sequence 15, Appl
36	1496	53.8	485	5	US-10-409-993-9 Sequence 9, Appl
37	1496	53.8	485	5	US-10-409-993-13 Sequence 13, Appl
38	1496	53.8	495	3	US-09-759-967-2 Sequence 2, Appl
39	1496	53.8	495	4	US-10-424-599-171705 Sequence 171705,
40	1496	53.8	495	4	US-10-425-114-45626 Sequence 45626, A
41	1496	53.8	501	4	US-10-425-114-43904 Sequence 43904, A
42	1496	53.8	505	4	US-10-425-114-71930 Sequence 71930, A
43	1496	53.8	505	4	US-10-425-114-71962 Sequence 71962, A
44	1496	53.8	506	4	US-10-425-114-43901 Sequence 43901, A
45	1496	53.8	506	4	US-10-425-114-45624 Sequence 45624, A

ALIGNMENTS

RESULT 1

US-10-245-871-96
; Sequence 96, Application US/10245871
; Publication No. US2003023594A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU, MINZHEN
; TITLE OF INVENTION: 11-KEY/ANTIGENIC EPIOTOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2013
; CURRENT APPLICATION NUMBER: US/10/245,871
; CURRENT FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-245-871-96

Alignment Scores:

Pred. No.: 1.95e-221 Length: 507
Score: 2673.00 Matches: 507
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 96.2% Indels: 0
DB: 4 Gaps: 0

US-10-728-323-3 (1-1524) x US-10-245-871-96 (1-507)

QY 1 CGCGAGCAACCGGAGGAGAACGCGTCCAGTTCAGCCCTCAATGCCAGACCTGAC 60

|||||

Db 1 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 20
QY 61 AATCGCATTAATCAGAGCGCGTTACATTGAGACTTGGAACCCCAACACAGAGAGTTC 120
Db 21 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 40
QY 121 GAATGCGCGCGTCCCTCTCTCGCTTAGTCTCCGCGCGCAACGCCCTTCGTAGGCTC 180
Db 41 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 60
QY 181 TTCTACTCCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGGGATACTTTGGTTG 240
Db 61 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 80
QY 241 ATATTCCCTCGTTGCTTAGACACTATGAAGAGCCTTCACACAAGGTGCTGCATCTCAG 300
Db 81 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgSerGln 100
QY 301 TCCCAAGACCAACAGAGCTTCCAAAGGAGAGAGCAACCAAGCAACAGAGATAGT 360
Db 101 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 120
QY 361 CACCAGAAGGTGCACCGTTTCAGATGAGGTGATCTCATTCAGTTCACCAGGTGTGCT 420
Db 121 HisGlnIleValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 140
QY 421 TTCTGCTCTPACCAACGACCACTGATGTGTTGCTGTTCTTCTTACTGACACCAAC 480
Db 141 PheTrpLeuTyrAsnAspHisAspThrAspValAlaValAsnLeuThrAspThrAsn 160
QY 481 AACCAACGACACAGCTTGATCAGTTCCTCCAGAGAGATTCNAATTTGCTGGACAGGAG 540
Db 161 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 180
QY 541 CAAGAGTTCCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGACTTACCATAT 600
Db 181 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 200
QY 601 AGCCCATACAGCCCGCAAGTTCAGCTTAGACAAGAAGAGCGTGAATTTAGCCCTCAGGA 660
Db 201 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 220
QY 661 CACACAGCCGAGAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 720
Db 221 GlnHisSerArgArgGluArgAlaGlyGlnGlnGluGluAsnGluGlyGlyAsnIlePhe 240
QY 721 AGCGGCTTACGCGGAGTTCCTGGAAACAGCTTCAGGTTCCAGCAGCAGCAGCAGCAGTAGT 780
Db 241 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 260
QY 781 CAAACCTAAGAGCGCAGACCGAGAGTGAAGAAGAGGAGCCATTGTGACAGTGAGGGGA 840
Db 261 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 280
QY 841 GCCTCAGATCTTGACCCAGATAGAAAGACGCTGCCAGCAAGAGAGGAAATACGAT 900
Db 281 GlyLeuArgIleLeuSerProAspArgLysArgAlaAspGluGluGluGluTyrAsp 300
QY 901 GAAGATCAATATGAATCAGATGAAGAGCAGTAGAAGCGTGGCAGGGGAAACAGAGCAGG 960
Db 301 GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg 320
QY 961 GGGAAATGGTATTGAAGAGAGATCTGCACCGCAGTCTTAAAGAACATTTGTTAGAAC 1020
Db 321 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 340
QY 1021 AGATCCCTGCATCTACACCTTCAGCTGTTCTCACTCAAACTGCAACAGCATCTCAAC 1080
Db 341 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 360
QY 1081 CTTCCTAATACTAGGTGGCTGGACCTTAGTGTGAATATGAATCTCTACAGGAATGCA 1140
Db 361 LeuLeuIleLeuArgTyrLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 380

QY 1141 TTGTTTTCGCTCACTACAACACCAACCGCACACATCATATCGATTGAGGGACGG 1200
Db 381 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 400
QY 1201 GCTCAGCTGCAAGTCTGTGACAGCAACCGCACAGAGTGTACGACGAGAGGCTTCAAGAG 1260
Db 401 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 420
QY 1261 GGTACGCTGCTTGTGCTCCACAGAACTTCGCCGCTCGCTGGAAAGTTCACAGACGAGAAC 1320
Db 421 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 440
QY 1321 TTCGAATAGCTGCATTCACAGACACTCAAGCCCGCAGCATAGCCAACTCCGCCGTGAA 1380
Db 441 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 460
QY 1381 AACTCCGTCATAGATAACCTGCCGAGAGGTGTGCAAAATTCATATGCGCTCCAAAGG 1440
Db 461 AsnSerValIleAspAsnLeuProGluGluValAlaAlaAsnSerTyrGlyLeuGlnArg 480
QY 1441 GAGCAGGCAAGCAGCTTAAGAACCAACCCCTTCAAGTTCCTCGTCCACCGTCTCAG 1500
Db 481 GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPheValProProSerGln 500
QY 1501 CAGTCTCCGAGGCTGTGCT 1521
Db 501 GlnSerProArgAlaValAla 507
RESULT 2
US-10-253-286-96
; Sequence 96, Application US/10253286
; Publication No. US20040058881A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; TITLE OF INVENTION: Ii-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2015
; CURRENT APPLICATION NUMBER: US/10/253,286
; PRIOR FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 905
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 96
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-253-286-96
Alignment Scores:
Pred. No.: 1,95e-221 Length: 507
Score: 2673.00 Matches: 507
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 96.2% Indels: 0
DB: 4 Gaps: 0
US-10-728-323-3 (1-1524) x US-10-253-286-96 (1-507)
QY 1 CGGCACAACCGGAGAGAGCGTCCAGTTCACGCGCTCAATGCGCAGACACCTGAC 60
Db 1 ArgGlnGlnProGlnGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 20
QY 61 AATCGCATTAATCAGAGCGCGTTACATTGAGACTTGGAACCCCAACACAGAGAGTTC 120
Db 21 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 40
QY 121 GAATGCGCGCGTCCCTCTCTCGCTTAGTCTCCGCGCGCAACGCCCTTCGTAGGCTC 180
Db 41 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 60

QY 181 TTCTACTCCAAATGCTCCCGAGAGATCTTATCCAGCAAGGAGCGGATACCTTTGGGTG 240
Db 61 PheTyrSerAsnAlaProGlnGluLeuPheIleGlnGlnGlyArgGlyTyrPheGlyLeu 80
QY 241 ATATTCCCTGGTGTCTTAGACATATGAAGAGCTCTACACAAAGGTCGTCGATCTCAG 300
Db 81 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlnGlyArgSerGln 100
QY 301 TCCCAAGACCAACCAAGAGCTCTCAAGGAGAGAGACCAAGCCACACACGACGATAGT 360
Db 101 SerGlnArgProProArgArgLeuGlnGlyAspGlnSerGlnGlnGlnArgAspSer 120
QY 361 CACCAAGAAGTCCACCGTTTCATCAGGCTGATCTCATTCGAGTTCCACCGGTGTGCT 420
Db 121 HisGlnIleValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 140
QY 421 TTCTGGCTCTACAACGACACGACTGATGTGTGTGTGTCTTCTTACTGACACCAAC 480
Db 141 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 160
QY 481 AACACGACACCAAGCTTGATCAGTCTCCAGGAGATTCATTTGGCTGGGACACGAG 540
Db 161 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 180
QY 541 CAAGAGTTCTTAAGGTACCAGCAACAAAGCAGACAAAGCAGACGAAGCTTACCATAT 600
Db 181 GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgSerLeuProTyr 200
QY 601 AGCCCATACAGCCCGCAAGTCCAGCTAGACAAAGAGCGGTGAATTTAGCCCTCAGGA 660
Db 201 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 220
QY 661 CAGCACACCCGACAGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 720
Db 221 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluAsnGluGlyGlyAsnIlePhe 240
QY 721 AGCGGCTTCACCGCGGAGTTCCTGCAACAAAGCTTCCAGTTGACGACGACGATAGT 780
Db 241 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 260
QY 781 CAAACCTTAAGAGGCGAGACCGAGAGTGAAGAGAGGAGCGCATTTGACAGTGAGGGA 840
Db 261 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 280
QY 841 GCGCTCAGATCTTCAGCCCATAGAAAGACGCTGCCGACGACGACGACGACGACGAC 900
Db 281 GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluGluTyrAsp 300
QY 901 GAAGATGAATATGAATACGATCAAGAGGATAGAAGCGTGGCAGGCGGAGACGAGCAGG 960
Db 301 GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg 320
QY 961 GCGAATGTGATTGAAGAGACGATCTGCACCGCAAGTGTCTTAAAGAACATTTGGTAGAAC 1020
Db 321 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysAsnIleGlyArgAsn 340
QY 1021 AGATCCCTTGACATCTAGACACCTCAAGCTGTTTCACTCAAACTGCCACGATCTCAAC 1080
Db 341 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 360
QY 1081 CTTCTTAATCTTAGTGGCTTCGACCTAGTGTGTAATATGGAATCTCTACAGGAATGCA 1140
Db 361 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 380
QY 1141 TTGTTTGTTCGCTCACTACAAACCAACGACGACGACGACGACGACGACGACGACG 1200
Db 381 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 400
QY 1201 GCTCAGTCAAGTGTGTGACGACGACGACGACGACGACGACGACGACGACGACGACG 1260
Db 401 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 420

QY 1261 GGTCACTGCTGTGTGTGTCCTCCACAGAACTTCGCCGTGCTGTGAAAGTCCCAGAGCGAGAAC 1320
Db 421 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 440
QY 1321 TTCGAATACGTGGCATTCACAGACAGACTCAAGGCCAGCATAGCCAACTTCGCCGTGAA 1380
Db 441 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 460
QY 1381 AACTCCGTCTAGATAAATCTCCGAGGAGGTGTGCAAAATTCATATATGCGCTCCAAAGG 1440
Db 461 AsnSerValIleAspAsnLeuProGluValValAlaAsnSerTyrGlyLeuGlnArg 480
QY 1441 GAGCAGGCAAGCAGCTTAAAGAACAAACCCCTTCAAGTTCTTCGTTCCACCGCTCAG 1500
Db 481 GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProProSerGln 500
QY 1501 CAGTCTCCGAGGCTGTGGCT 1521
Db 501 GlnSerProArgAlaValAla 507
RESULT 3
US-10-228-806-6
; Sequence 6, Application US/10228806
; Publication No. US20030049237A1
; GENERAL INFORMATION:
; APPLICANT: Bannion, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reactions
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0043
; CURRENT APPLICATION NUMBER: US/10/228,806
; CURRENT FILING DATE: 2002-08-26
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-10-228-806-6
Alignment Scores:
Pred. No.: 1,95e-221 Length: 510
Score: 2673.00 Matches: 507
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 96.2% Indels: 0
DB: 4 Gaps: 0
US-10-728-323-3 (1-1524) x US-10-228-806-6 (1-510)
QY 1 CGGCAGCAACCGGAGGAGAACCGTGCAGTTCCAGCGCTCAATGCCGAGACCTGAC 60
Db 4 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 23
QY 61 AATCGATTGAATCAGAGGGCGGTTCATTGAGACTTGGAAACCCCAACACAGAGATTTC 120
Db 24 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 43
QY 121 GAATCGCGCGCGTCCCTCTCGCTTAGTCTCTCCGCGCAACGCCCTTCGTAGGCCT 180
Db 44 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 63
QY 181 TTCTACTCCAATGTCTCCCGAGGAGATCTTCATCCAGCAAGGAGGAGGATCTTTGGGTG 240
Db 64 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 83
QY 241 ATATTCCCTGGTGTCTTAGACATATGAAGAGCTCTACACAAAGGTCGTCGATCTCAG 300
Db 84 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlnGlyArgSerGln 103
QY 301 TCCCAAGACCAACCAAGAGCTCTCCAGGAGAGACCAAGCCACACAGCAAGATAGT 360
Db 104 SerGlnArgProProArgArgLeuGlnGlyAspGlnSerGlnGlnGlnArgAspSer 123

QY 361 CACCAGAGGTGCACCGTTTCGATGAGGGTGATCTCATTCGAGTTCCACCGGTGTGCT 420
DB 124 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143
QY 421 TTCTGGCTCTACAAACACACGACACTGATGTTGTTGCTGTTCTCTTACTGACACCAAC 480
DB 144 PheTrpLeuTyrrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 163
QY 481 AACACGACAACAGCTTGATCAGTTCCCGAGAGATTCAATTTGGCTGGGACACGGAG 540
DB 164 AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnThrGlu 183
QY 541 CAAGAGTTCTTAGGTACAGACACAAACGACAAAGCAGACGACGAGAGAGCTTACCATAT 600
DB 184 GlnGluPheLeuArgTyrrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyrr 203
QY 601 AGCCCATACAGCCCGCAAGTTCAGCTTAGACAAAGAGCGTGAATTTAGCCCTCGAGGA 660
DB 204 SerProTyrrSerProGlnSerGlnProArgGlnGlnGluArgGluPheSerProArgGly 223
QY 661 CAGCACGCGCAGACAGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 720
DB 224 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluAsnGluGlyAsnIlePhe 243
QY 721 AGCGGCTTCACGCGGAGTTCCTGGAAACAGCCCTTCAGGTTCCAGCACGACGACGACG 780
DB 244 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 263
QY 781 CAAACCTAAGAGCGCAGACCGAGAGTGAAGAGAGAGGAGCCATTGTGACAGTGAGGGGA 840
DB 264 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 283
QY 841 GGCCTCAGAACTTGAGCCCGACAGATAGAAAGAGACGTGCCGACGAAAGAGGAATACGAT 900
DB 284 GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrrAsp 303
QY 901 GAGATCAATATGAATACGATGAAGAGATAGAGCGGTGGCAGGGGACGACGAGCAGG 960
DB 304 GluAspGluTyrrGluTyrrAspGluGluAspArgArgArgGlyArgGlySerArgGlyArg 323
QY 961 GCGAATGGTATTGAAGACGATCTGCACCGCAAGTGCTAAAAAGAACATTTGTTAGAAAC 1020
DB 324 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 343
QY 1021 AGATCCCTCGACATCTACAAACCTCAAGCTGGTTCACTCAAACTGCGCAACGATCTCAAC 1080
DB 344 ArgSerProAspIleTyrrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 363
QY 1081 CTTCTAATACTTAGTGGCTTGACCTAGTGTGTAATATGGAATCTTACAGGAATGCA 1140
DB 364 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrrGlyAsnLeuTyrrArgAsnAla 383
QY 1141 TTGTTTGTCTCACTACAAACACCGCACACGACATCATATATCATTTGAGGGGACGG 1200
DB 384 LeuPheValAlaHisTyrrAsnThrAsnAlaHisSerIleIleTyrrArgLeuArgGlyArg 403
QY 1201 GCTCAGCTGCAAGTGTGGGACAGCAACGCAACAGAGTGACACGAGGAGCTTCAAGAG 1260
DB 404 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrrAspGluGluLeuGlnGlu 423
QY 1261 GGTCAAGTCTGTGGTGCCACAGACTTCGCGCTGCGTGAAGTCCCGAGCGGAGAAC 1320
DB 424 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 443
QY 1321 TTGGAATACGTGGCATTCAGACAGACTCAAGGCCCGACATAGCCAACTTCGCGGTGAA 1380
DB 444 PheGluTyrrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 463
QY 1381 AACTCCGTCATAGATAACCTGCGGAGGAGGTGGTTGCAATTCATATGAGCTCCAAAGG 1440
DB 464 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrrGlyLeuGlnArg 483
QY 1441 GAGCAGGCAAGGCAGCTTAAGAACAAACACCCCTTCAAGTTCTTCGTTCCACCGTCTCAG 1500

DB 484 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProProSerGln 503
QY 1501 CAGTCTCCGAGGCTGTGGCT 1521
DB 504 GlnSerProArgAlaValAla 510
RESULT 4
US-10-100-303A-90
; Sequence 90, Application US/10100303A
; Publication No. US20030202980A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, et al.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction
; TITLE OF INVENTION: to Allergy
; FILE REFERENCE: 2002834-0166
; CURRENT APPLICATION NUMBER: US/10/100,303A
; CURRENT FILING DATE: 2002-03-18
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 90
; TYPE: PRT
; LENGTH: 510
; ORGANISM: Atachis hypogaea, Prot/Nucleo Ara h 3
US-10-100-303A-90
Alignment Scores:
Pred. No.: 1,95e-221 Length: 510
Score: 2673.00 Matches: 507
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 96.2% Indels: 0
DB: 4 Gaps: 0
US-10-728-323-3 (1-1524) x US-10-100-303A-90 (1-510)
QY 1 CGSCAGCAACCGGAGAGAACGCGTCAGGTCAGGCTCAATGCGGAGACGCTGAC 60
DB 4 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 23
QY 61 AATCGCATTTGAATCAGAGCGGTTACATTGAGACTTGGAACCCCAACACAGGAGTTC 120
DB 24 AsnArgIleGluSerGluGlyGlyTyrrIleGluThrTrpAsnProAsnAsnGlnGluPhe 43
QY 121 GAATGCGCGGCGTCCCGCTCTCTCGCTTAGTCTCCTCGCGGCAACGCCCTTCCTAGCCCT 180
DB 44 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 63
QY 181 TTCTACTCCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGATACTTTGGGTTG 240
DB 64 PheTyrrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrrPheGlyLeu 83
QY 241 ATATTCTCTGGTTGTCCTAGACATATGAAGAGCCTCACACAAAGTCGTCGATCTCAG 300
DB 84 IlePheProGlyCysProArgHisTyrrGluGluProHisThrGlnGlyArgArgSerGln 103
QY 301 TCCCAAGACCAACAGACGCTCTCCAGGAGAGACCAAGCCCAACAGCAACAGAGTAGT 360
DB 104 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 123
QY 361 CACCAGAAGGTGCACCGTTTCGATGAGGGTGATCTCATTCGAGTTCCCGCGGTGTGCT 420
DB 124 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 143
QY 421 TTCTGGCTCTACAAACACCAACGACACTGATGTTGTTGCTGTTCTCTTACTGACACCAAC 480
DB 144 PheTrpLeuTyrrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 163
QY 481 AACCAACGACAACAGCTTGATCAGTTCCCGCAGAGATTCGAATTTGGCTGGGACACCGGAG 540
DB 164 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 183
QY 541 CAAGAGTTCTTAAGGTACCGACCAACAAAGCAGACAAAGCAGACGAGAGAGCTTACCATAT 600

Db 184 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 203
QY 601 AGCCATACAGCCGCAAGTTCAGCTAGACAAAGAGCGTGAATTTAGCCCTCGAGA 660
Db 204 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgPheSerProArgGly 223
QY 661 CAGCACAGCCGAGAGAACGAGCAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 224 GlnHisSerArgArgGluArgAlaGlyGlnGlnGluGluGluGluGluGluGluGlu 243
QY 721 AGCGCTTACCCCGAGTTCCTGGACCAAGCTTCCAGCTTCCAGCTTCCAGCAGCAG 780
Db 244 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 263
QY 781 CAAACCTTAAGAGCCGAGACCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 264 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGluGluGluGluGluGluGlu 283
QY 841 GGCCTCAGAAATCTTGAGCCAGATAGAAAGAGACGTCGCGACGAGAGAGAGAGAG 900
Db 284 GlyLeuArgIleLeuSerProAspArgLysArgAlaAspGluGluGluGluGluGlu 303
QY 901 GAAGATCAATATCATACATCAAGAGGATAGAGCGCTGGCAGGGAGAGCAGGCAG 960
Db 304 GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg 323
QY 961 GCGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGTAAAGAGAGAGAGAGAGAG 1020
Db 324 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArg 343
QY 1021 AGATCCCTTGACATCTACAACTCAAGCTTCACTCAAACTGCCAAGCATCTCAAC 1080
Db 344 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeu 363
QY 1081 CTTCTAATCTTAGTGGCTTGGACCTAGCTGTGTAATATGGAATCTTACAGGATGCA 1140
Db 364 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 383
QY 1141 TTGTTTGTCTGCTCCTACCAACCAACGACACAGCATCATATATCGATTGAGGGGACG 1200
Db 384 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 403
QY 1201 GTCACGTCAAGTCGTGACAGCAACCGCAACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
Db 404 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 423
QY 1261 GGTACGCTCTGTGTGTCACAGAACTTCGCGCTGCTGGAAAGTCCAGAGCGAGAAC 1320
Db 424 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 443
QY 1321 TTCGAATACGTGCTTCAAGACAGACTCAAGGCCAGCATAGCCCAACCTCCCGGTGAA 1380
Db 444 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 463
QY 1381 AACTCCGTATAGATAAATCCGCGAGGAGGTGGTTCGAAATTCATATATGCTCCCAAGG 1440
Db 464 AsnSerValIleAspAsnLeuProGluGluValAlaAsnSerTyrGlyLeuGlnArg 483
QY 1441 GAGCAGGCAAGGAGCTTAAGAACCAACACCCCTTCAAGTCTTCTGTTCCACCGTCTAG 1500
Db 484 GluGlnAlaArgGlnLeuLysAsnAsnProPheLysPhePheValProProSerGln 503
QY 1501 CAGTCTCCGAGGCTGTGGCT 1521
Db 504 GlnSerProArgAlaValAla 510

RESULT 5
US-09-731-221-79
; Sequence 79, Application US/09731221
; Patent No. US20020018778A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael

; TITLE OF INVENTION: Passive Desensitization
; FILE REFERENCE: 2002834-0103
; CURRENT APPLICATION NUMBER: US/09/731,221
; CURRENT FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Arachis
; OTHER INFORMATION: Hypogaea
US-09-731-221-79

Alignment Scores:
Pred. No.: 2,07e-218 Length: 526
Score: 2638.00 Matches: 502
Percent Similarity: 99.0% Conservative: 0
Best Local Similarity: 99.0% Mismatches: 5
Query Match: 94.9% Indels: 0
DB: 3 Gaps: 0

US-10-728-323-3 (1-1524) x US-09-731-221-79 (1-526)
QY 1 CGGCAGCAACCGAGGAGAGACGCTGCCAGTTCAGCGCTCAATGCCGAGAGACCTGAC 60
Db 5 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 24
QY 61 AATCGCATTTGAATCAGAGGGGGTTTACATTGAGACTTGGAAACCCCAACACCCAGAGTTC 120
Db 25 AsnArgIleGluSerGluGlyGlyTyrIleGluThrTrpAsnAlaAsnGlnGluPhe 44
QY 121 GAATCGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCGCTTCGTAGGCCT 180
Db 45 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 64
QY 181 TTCTACTCCAATGTCTCCAGGAGATCTTCATCCAGCAAGGAGGGGATCTTTGGGTG 240
Db 65 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnArgGlyTyrPheGlyLeu 84
QY 241 ATATTCCCTGTTGCTCTAGACACTATGAAGAGCTTCACACAGGCTCGTGCATCTCAG 300
Db 85 IlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln 104
QY 301 TCCCAAGACACCAAGACGCTCTCAAGCAGAGAACCAAGCAACAGCAACAGATAGT 360
Db 105 SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnArgAspSer 124
QY 361 CACCAAGAGTGCACCGTTCGATGAGGGTGTCTTCATTGCGAGTTCCACCGGTGTGCT 420
Db 125 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 144
QY 421 TTCTGGCTCTACACGACACGACACTGATGTTGTGTGTTCTTCTTACTGACACCAAC 480
Db 145 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 164
QY 481 AACCAACCAACCGCTTGTATCAGTTCCTCCAGGAGATTCAATTCGCTGGGAAACCGAG 540
Db 165 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu 184
QY 541 CAAAGAGTCTTAAAGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCATAT 600
Db 185 GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgArgSerLeuProTyr 204
QY 601 AGCCATACAGCCCGCAAGTTCAGCTAGACAAAGAGAGCGTGAATTTAGCCCTCGAGA 660
Db 205 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 224
QY 661 CAGCACCGCCGAGAGACGAGCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 225 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluGluGluGluGluGluGlu 244

```
QY 721 AGCGGCTTCACGCCGAGTTCCTGGAAACAAGCCTCCAGGTGACACACAGACAGATAGTG 780
Db 245 SerGlyPheThrProGluAlaLeuPheGlnAlaPheGlnValAspArgGlnIleVal 264
QY 781 CAAACCTTAAGAGCCAGACCCAGAGTGAAGAGAGGAGCCATTGTGACAGTGAGGGGA 840
Db 265 GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly 284
QY 841 GGCTCAGAAATCTTGAGCCAGATAGAAAGAGACGTGCCGACGAAGAAGAGGAAATACGAT 900
Db 285 GlyLeuArgAlaLeuSerProAspArgLysArgAlaAspGluGluGluThrAsp 304
QY 901 GAAGATGAATATCAATACCATGAAGAGATAGAAAGCGTGGCAGGGGAAGCAGAGCAGG 960
Db 305 GluAspGluThrAlaThrAspGluGluAspArgArgGlyArgGlySerArgGlyArg 324
QY 961 GGAATCGTATTGAAGAGACGATCTGCACCGCAAGTGTAAAGAAACATTTGTAGAAAC 1020
Db 325 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 344
QY 1021 AGATCCCTGACATCTACAACCCCTCAAGCTGGTTCATCAAACTGCCAAGCATCTCAAC 1080
Db 345 ArgSerProAspIleThrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 364
QY 1081 CTTCTAATCTTAGGTGGCTTGACCTAGTGTGAATATGGAATCTCTACAGGAATGCA 1140
Db 365 LeuLeuIleuLeuArgIleuGlyProSerAlaGluThrGlyAsnLeuThrArgAsnAla 384
QY 1141 TTCTTTCTGCTCACTACAACACCAACCGCACACAGCATCATATATCGATTGAGGGGACGG 1200
Db 385 LeuPheValAlaHisThrAsnThrAsnAlaHisSerIleIleIleThrArgLeuArgGlyArg 404
QY 1201 GCTCAGTGAAGTCTGTGGACACAAACGGCAAACAGAGTGTACACAGAGAGCTTCAAGAG 1260
Db 405 AlaHisValGlnValValAspSerAsnGlyAsnArgValThrAspGluGluLeuGlnGlu 424
QY 1261 GGTCAAGTCTGTGGTCCACAGAACTTCGCGCTCGTGGAAAGTCCACAGACGAGAAC 1320
Db 425 GlyHisValLeuValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 444
QY 1321 TTCGAATACGTGCATTCAAGACAGACTCAAGGCCCCAGCATAGCCAACTCGCGGTGAA 1380
Db 445 PheGluThrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 464
QY 1381 AACTCGTCTATAGATAACCTGCCGAGAGAGGTGTGCAAAATTCATATGCCCTCCAAAGG 1440
Db 465 AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerThrGlyLeuGlnArg 484
QY 1441 GACGAGCAAGGCGAGCTTAAGAACAACACCCCTTCAAGTTCTTCTGTTCCACCGTCTCAG 1500
Db 485 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProProSerGln 504
QY 1501 CAGTCTCCGAGGCGTGTGGCT 1521
Db 505 GlnSerProArgAlaValAla 511
RESULT 6
US-10-899-551-6
; Sequence 6, Application US/10899551
; Publication No. US2005063994A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Burks, A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Howard, Sosin B.
; APPLICANT: Bottomly, Kim H.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
; FILE REFERENCE: 2002834-0233
; CURRENT APPLICATION NUMBER: US/10/899,551
; CURRENT FILING DATE: 2004-07-26
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
```

```
; LENGTH: 530
; TYPE: PRT
; ORGANISM: species Arachis hypogea
US-10-899-551-6
Alignment Scores:
Pred. No.: 6,08e-217 Length: 530
Score: 2621.00 Matches: 500
Percent Similarity: 98.6% Conservative: 0
Best Local Similarity: 98.6% Mismatches: 7
Query Match: 94.3% Indels: 0
DB: Gaps: 0
US-10-728-323-3 (1-1524) x US-10-899-551-6 (1-530)
QY 1 CGGAGCAACCGAGAGAGAACCGGTGCCAGTTCAGCGCTCAATCGCGCAGACACTGAC 60
Db 24 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
QY 61 AATCGCATTGAATCAGAGGGCGGTTACATTGAGACTTGGAAACCCCAACCAACAGGAGTTTC 120
Db 44 AsnArgIleGluSerGluGlyThrIleGluThrTrpAsnProAsnAsnGlnGluPhe 63
QY 121 GAATGCGCGGCGTCCGCTCTCTCGCTTAGTCTCTCCGCCGCAACGCCCTTCGTAGCCCT 180
Db 64 GluCysAlaGlyValAlaAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 83
QY 181 TTCTACTCCAATCTCCCCAGAGAGATCTTCATCCAGCAAGGAGGGGATACTTTGGTTG 240
Db 84 PheThrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyThrPheGlyLeu 103
QY 241 ATATTCCCTGGTGTCTCTAGACACATATGAAGAGCCTCACACACAAGTTCGTGATCTCAG 300
Db 104 IlePheProGlyCysProSerThrThrGluGluProAlaGlnGlnGlyArgArgSerGln 123
QY 301 TCCCAAGACCAACCAAGACGTCTCCAAGGAGAGAACCAAAAGCCACACAGCAACAGATAGT 360
Db 124 SerGlnArgProIleArgLeuGlnGlyGluAspGlnSerGlnGlnGlnGlnArgAspSer 143
QY 361 CACCAAGAGTGCACCGTTTCATGAGGTGTATCTCATTCAGTTCCTCCACCGGTGTGCT 420
Db 144 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 163
QY 421 TTCTGGCTCTACAACACCAACACGACACTGATGTTGTTGCTGTTCTCTCTACTGACACCAAC 480
Db 164 PheThrLeuThrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 183
QY 481 AACCAACGACCAACCGCTTGATCAGTTCCCGAGAGATTCAATTGCTGGGAACACGGAG 540
Db 184 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu 203
QY 541 CAAGAGTTCTTAAGTTACAGCAACAACAGCAGCAACAGCAGAGAGAGAGCTTACCATAT 600
Db 204 GlnGluPheLeuArgThrGlnGlnGlnSerArgGlnSerArgArgArgSerLeuProThr 223
QY 601 AGCCCATACAGCCCGCAAAAGTACAGCTAGACAAGAAGAGCGTCAATTATTCCTCCAGGA 660
Db 224 SerProThrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 243
QY 661 CAGCACAGCCGACAGAGACGAGCAGGACAAGAGAAGAAAGAAACCAAGGTGGAAACATCTTC 720
Db 244 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluGluGluGlyGlyAsnIlePhe 263
QY 721 AGCGGCTTCACGCCGAGTTCCTGGAAACAAGCCCTCCAGGTTCACACACAGACAGATAGT 780
Db 264 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 283
QY 781 CAAACCTTAAGAGCCAGACCCAGAGTGAAGAGAGGAGCCATTGTGACAGTGAGGGGA 840
Db 284 GlnAsnLeuArgGlyGluAsnGluSerGluGluGluGlyAlaIleValThrValArgGly 303
QY 841 GGCCTCAGAAATCTTGAGCCCGATAGAAAGAGACGTGCCGACGAGAGAGGAATACGAT 900
```

```
Db 304 GlyLeuArgIleLeuSerProAspArgLysArgAlaAspGluGluGluTyrAsp 323
QY 901 GAAGATGAATATCAATACCATGAAGAGGATAGAAGCGTGGCAGGGAGAGCAGG 960
Db 324 GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlyArgGlyArg 343
QY 961 GGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGTAAAGAACATTCGTGAAC 1020
Db 344 GlyAsnGlyIleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn 363
QY 1021 AGATCCCTCAGATCTACCAACCTCAGCTGCTCACTCAAACTGCCAACCATCTCAAC 1080
Db 364 ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn 383
QY 1081 CTTCTAATACTTAGTGGCTTGACCTAGTGTGTAATATGGAATCTCTACAGGAATGCA 1140
Db 384 LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla 403
QY 1141 TTGTTTGTCTCCTACCAACCAACCGCACAGCATCATATATCGATTGAGGGACGG 1200
Db 404 LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg 423
QY 1201 GCTCAGCTCGAAGTCTGACAGCAACCGCACAGAGTGTACGACGAGGAGCTTCAAGAG 1260
Db 424 AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu 443
QY 1261 GGTACAGCTGTTGTGTGTCACAGAACTTCGCGCTCGTGGAAAGTCCACAGACGAGAAC 1320
Db 444 GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn 463
QY 1321 TTGGAATACGTGGCATTCAGACAGACTCAAGGCCAGCATAGCCAACTCCCGGTGAA 1380
Db 464 PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu 483
QY 1381 AACTCGCTCATAGATAACTCGCGAGGAGGTGTTGCAAAATCATATGCCCTCCAAAGG 1440
Db 484 AsnSerValIleAspAsnLeuProGluGluValAlaAlaAsnSerTyrGlyLeuGlnArg 503
QY 1441 GAGCAGGACGAGGAGCTTAAGAACAAACACCCCTTCAAGTCTTCTGTTCCACCGTCTCAG 1500
Db 504 GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProProSerGln 523
QY 1501 CAGTCTCCAGGCGTGTGGCT 1521
Db 524 GlnSerProArgAlaValAla 530

RESULT 7
; Sequence 58, Application US/10899551
; Publication NO. US20050063994A1
; GENERAL INFORMATION:
; APPLICANT: Caplan, Michael J.
; APPLICANT: Burks, A. Wesley
; APPLICANT: Sampson, Hugh A.
; APPLICANT: Howard, Sosin B.
; APPLICANT: Bottomly, Kim H.
; TITLE OF INVENTION: Methods and Reagents for Decreasing Clinical Reaction to Allergy
; FILE REFERENCE: 2002834-0233
; CURRENT APPLICATION NUMBER: US/10/899,551
; CURRENT FILING DATE: 2004-07-26
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 58
; LENGTH: 351
; TYPE: PRT
; ORGANISM: species Arachis hypogea
US-10-899-551-58

Alignment Scores:
Pred. No.: 1 63e-134 Length: 351
Score: 1665.00 Matches: 315
Percent Similarity: 97.8% Conservative: 0
Best Local Similarity: 97.8% Mismatches: 7
```

```
Query Match: 59.9% Indels: 0
DB: 5 Gaps: 0
US-10-728-323-3 (1-1524) x US-10-899-551-58 (1-351)
QY 1 CGGACAGCAACCGGAGAGAAACCGTGCAGTTCACAGCCCTCAATGCGCAGACAGCTGAC 60
Db 15 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 34
QY 61 AATCCGATTGAATCAGAGGGCGTTACATTGAGACTTGGAACCCCAACCAACAGGAGTTC 120
Db 35 AsnArgIleGluSerGluGlyGlyIleGluThrTrpAsnProAsnAsnGlnGluPhe 54
QY 121 GAATCGCGCGCGTCCGCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCCT 180
Db 55 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 74
QY 181 TTCTACTCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGGATACTTTGGTTG 240
Db 75 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 94
QY 241 ATATTCCCTGTTGTCTAGACACTATCAAGAGCCTCACACAAGGTCTGTCGATCTCAG 300
Db 95 IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgArgSerGln 114
QY 301 TCCCAAGACCAACCAAGACGCTCTCAAGGAGAAGACCAAGCCAAACAGCAACGAGATAGT 360
Db 115 SerGlnArgProProIleArgLeuGlnGlyGluAspGlnSerGlnGlnGlnArgAspSer 134
QY 361 CACCAAGAGGTGCACCGTTTCGATGAGGGTGAATCTCATTCGACAGTTCCACCGGTGTGCT 420
Db 135 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 154
QY 421 TTCTGGCTCTACACGACGACGACTGATCTGTTGCTGTTCTTCTTACTGACACCCAC 480
Db 155 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValAsnLeuThrAspThrAsn 174
QY 481 AACACACCAACCAAGCTTGATCAGTTCCTCCAGGAGATTCATTTTCGCTGGGAACACGAG 540
Db 175 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu 194
QY 541 CAAGAGTTCTTAAGTACACGACCAACAAAGCAGACAAAGCAGACGAGAGCTTACCATAT 600
Db 195 GlnGluPheLeuArgTyrGlnGlnSerArgGlnSerArgArgArgSerLeuProTyr 214
QY 601 AGCCCATACAGCCCGCAAGTCAGCTAGACAGAGAGCGCTGAATTTAGCCCTCGAGA 660
Db 215 SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly 234
QY 661 CAGCACAGCCGACAGAAACGAGCAGACAGCAAGAAAGAAAGAAAGAGGTGGAAACATCTTC 720
Db 235 GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluAsnGluGlyAsnIlePhe 254
QY 721 AGCGGCTTCACGCGGAGTTCCTGTAACAAGCTTCAGGTTCCAGTTGACGACAGACAGTAGTG 780
Db 255 SerGlyPheThrProGluPheLeuGlnAlaPheGlnValAspAspArgGlnIleVal 274
QY 781 CAABACCTAAGAGCGGACCGCAGAGTGAAGAGAGGAGCGCATTTGACAGTGAAGGGA 840
Db 275 GlnAsnLeuArgGlyGluAsnGluSerGluGluGlyAlaIleValThrValArgGly 294
QY 841 GGCCTCAGAATCTTTGAGCCGACAGATAGAAGAGACGTCGCCGACGAAGAAGAGGAATACGAT 900
Db 295 GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluGluTyrAsp 314
QY 901 GAAGATGAATATGAATACGATGAAGAGGATAGAAGCGCTGCGAGGGAGACGAGCGAGG 960
Db 315 GluAspGluTyrGluTyrAspGluGluAspArgArgArgGlyArgGlySerArgGlyArg 334
QY 961 GGGAT 966
Db 335 GlyAsn 336
```

RESULT 8

```

US-09-759-967-21
; Sequence 21, Application US/09759967
; Publication No. US20030166518A1
; GENERAL INFORMATION:
; APPLICANT: The Board of Regents of the University of Nebraska
; TITLE OF INVENTION: METHOD FOR ALLERGEN CHARACTERIZATION
; FILE REFERENCE: UNL 3001.01
; CURRENT APPLICATION NUMBER: US/09/759,967
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Glycine max
US-09-759-967-21

Alignment Scores:
Pred. No.: 2,12e-120 Length: 481
Score: 1502.00 Matches: 297
Percent Similarity: 71.6% Conservative: 68
Best Local Similarity: 58.2% Mismatches: 91
Query Match: 54.0% Indels: 54
DB: 3 Gaps: 7

US-10-728-323-3 (1-1524) x US-09-759-967-21 (1-481)

QY 1 CGGCAGCAACCGGAGGAGAAACGCGTCCAGCTTCCAGCGCCCTCAATGCGCAGAGACCTGAC 60
DB 23 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 42
QY 61 AATGCGATGAATCAGAGGGCGGTTCATTGAGACTTGGAACCCCAACACGAGAGTTC 120
DB 43 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 62
QY 121 GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCT 180
DB 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 82
QY 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGGATCTTTGGTTG 240
DB 83 SerTyrThrAsnAlaProGlnGlnIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 102
QY 241 ATATTCCCTGGTTGCTTAGACACTATGAAGAGCCTCACACAAAGTTCGTGATCTCAG 300
DB 103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 119
QY 301 TCCCAAGACCAACCAAGAGCTCTCCAAGGAGAGAACCAAGCCACGACGATAGT 360
DB 120 SerSerArgPro-----GlnAspArg 126
QY 361 CACCAAGAGGTGCACCGTTTCGATGAGGTGATCTCATTGCAGTTCCCAACCGGTGCT 420
DB 127 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146
QY 421 TTCGGCTCTCAACAGCACCAAGACACTGATGTGTGCTGTTCTCTTACTGACACCAAC 480
DB 147 TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166
QY 481 AACACAGCAACACGAGCTTGATCAGTTCCCGAGGAGATCAATTCGGCTGGGAACACGGAG 540
DB 167 SerPheGlnAsnGlnLeuAspGlnMetProAlaArgPheTyrLeuAlaGlyAsnGlnIle 186
QY 541 CAAGAGTCTTAAAGTTACAGCAACAAAGACGACAAAGCAGACGAAAGAGCTTACATAT 600
DB 187 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly----- 199
QY 601 AGCCCATACAGCCCGCAAGTCAAGCTCAGCCTAGACGAAGAGCGTGAAATTTAGCCCTCGAGA 660
DB 200 -----Gly 200

QY 661 CAGCACAGCCGAGAGAAACGACGAGCAACGAAGAAGAAAGAAACGAAGGTGGAAACATCTTC 720

```

```

; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 171702
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_126061C.1.pcp
; US-10-424-599-171702

```

Alignment Scores:

```

Pred. No.: 2,12e-120 Length: 481
Score: 1502.00 Matches: 297
Percent Similarity: 71.6% Conservative: 68
Best Local Similarity: 58.2% Mismatches: 91
Query Match: 54.0% Indels: 54
DB: Gaps: 7

```

US-10-728-323-3 (1-1524) x US-10-424-599-171702 (1-481)

```

QY 1 CGCAGCAACCGGAGGAGAACGGGTGCCAGTTCACAGCGCTCAATCGCGAGACACTGAC 60
Db 23 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 42
QY 61 AATCGCATTAATCAGAGCGGTTCATATGAGACTTGAACCCCAACCAACAGGAGTTC 120
Db 43 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 62
QY 121 GAATGCGCGGCTCGCCCTCTCTCGTGTAGTCTCCGCGCAACGCCCTTCGTAGCGCT 180
Db 63 GlnCysAlaIleValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 82
QY 181 TTCTACTCCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGAGATCTTTGGTTG 240
Db 83 SerTyThrAsnAlaProGlnGlnIleTyIleGlnGlnGlySerGlyIlePheGlyMet 102
QY 241 ATATTCCTCGTGTCTAGACACTATGAGAGCCTCACACAAAGTGTGTCATCTCAG 300
Db 103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 119
QY 301 TCCCAAGACCCCAAGACGCTCTCCAGGAGAGAACCAAGCCCAACAGCAACGAGATAGT 360
Db 120 SerSerArgPro-----GlnAspArg 126
QY 361 CACCAAGAGTGCACCGTTTCGATGAGGTGATCTCAATTCAGTTCACCGGTGTGCT 420
Db 127 HisGlnLysIleTyHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146
QY 421 TTCGTGCTCTAACACACGACGACACTGATGTTGTTGTTGTTCTTCTTACTGACCAAC 480
Db 147 TyrTrpMetTyTrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166
QY 481 AACCAACGACACGACGTTGATCAGTTCCTCCAGGAGATTCATTTTGGCTGGGACACGGAG 540
Db 167 SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyTrLeuAlaGlyAsnGlnGlu 186
QY 541 CAAGAGTTCCTTAAGGTACCAAGCAACAAAGACAGCAAGAGGAGGAGGAGGAGGAGGAG 600
Db 187 GlnGluPheLeuGlnTyTrGlnProGlnLysGlnGlnGly-----199
QY 601 AGCCCATACAGCCCGCAAGATGACGCTTAGCAACAGAGAGCGGTGAATTTAGCCCTCAGGA 660
Db 200 -----Gly 200
QY 661 CAGCAGACGCGCAGAGAACGAGCAGGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 201 ThrGlnSerGlnLysGlyLysArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 220
QY 721 AGCGGTTCACGCGGAGTTCCTCGAACAAGCCTTCAGGTTTCAGCAGACAGACAGATAGTG 780
Db 221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239
QY 781 CAANAACCTAAGAGCGCAGACCGAGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db :::::||||| ||| |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

```

```

Db 240 ArgLysLeuGlnGlyGluAsnGluGluGluGluGlyValThrValLysGly 259
QY 841 GGCCTCAGAAATCTTGAGCCCA-----GATAGAAAGAGACGTCGCCGACGAAGAGAG 891
Db 260 GlyLeuSerValIleSerProThrGluGluGlnGlnArgProGluGluGluGlu 279
QY 892 GAATACGATGAAGATGAATATGAAATACGATGAAGAGGATAGAAAGCGTGGCAGGGAAGC 951
Db 280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294
QY 952 AGAGCAGCGGGGAATGTTTGAAGAGAGGATCTGCACCGCAAGTCTTAAAGAACATTT 1011
Db 295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311
QY 1012 GGTAGAAACAGATCCCTGACATCTACAAACCTCAAGCTGAGTTCACCTCAAAACCTGCAAC 1071
Db 312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrAlaThr 331
QY 1072 GATCTCAACCTTCTTAATACTTAGTGGTGTGACCTAGTGTGAATATGAAATCTCTAC 1131
Db 332 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 351
QY 1132 AGGAATGCAATCTTTTGTCTGCTCCTACACACCAACGACACAGCATCATATCGATTG 1191
Db 352 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371
QY 1192 AGGGACGCGGCTCAGTGCAGTCTGGGACAGCAACCGGCAACAGAGTGTACGACGAGGAG 1251
Db 372 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 391
QY 1252 CTTCAAGAGGTCAGTGTCTGTGTGTCACAGAACTTCGCGTGTGCGTGGAAAGTCCCAG 1311
Db 392 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411
QY 1312 AGCGAGAACTTCGATAGTGTGTCATTCAGACAGACTCAAGCCGACGATAGCCACCTC 1371
Db 412 SerAspAsnPheGluTyTrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431
QY 1372 GCGGTGAAACCTCCGTATAGTAACTCCGCGAGGAGGTGTGCAAAATTCATATGCG 1431
Db 432 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 451
QY 1432 CTTCAAGAGGACGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1491
Db 452 LeuArgArgGlnAlaArgGlnValLysAsnAsnAsnAsnProPheSerPheLeuValPro 471
QY 1492 CCGTCTCAGCAGTCTCCGAGGCGCTGTGCT 1521
Db 472 ProLysGluSerGlnArgArgValAlaAla 481

```

RESULT 10

```

US-10-425-114-43865
; Sequence 43865, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 43865
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700645818_FLI.pcp

```

US-10-425-114-43865

Alignment Scores:

Pred. No.: 2,13e-120 Length: 488
Score: 1502.00 Matches: 297
Percent Similarity: 71.6% Conservatives: 68
Best Local Similarity: 58.2% Mismatches: 91
Query Match: 54.0% Indels: 54
DB: 4 Gaps: 7

US-10-728-323-3 (1-1524) x US-10-425-114-43865 (1-488)

```
QY 1 CGGCAGCAACCGGAGAGAACGGTGTCCAGCTCAATCGGCAGACCTGAC 60
DB 30 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 49
QY 61 AATCGCATTAATCAGAGGGCGGTTACATTGACACTTGGACCCCAACACAGGAGTTTC 120
DB 50 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 69
QY 121 GAATGCGCGGCTCGCCCTCTCTCGCTTAGTCTCGCGCAACGCCCTTCGTAGCCCT 180
DB 70 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 89
QY 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGGGATACTTTGGTTG 240
DB 90 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 109
QY 241 ATATTCCCTGGTTGCTTAGACACTATGAAGAGCCTCACACAAAGTCTCGATCTCAG 300
DB 110 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 126
QY 301 TCCCAAGACCAACCAAGACGTCTCCAAGGAGAGACCAAGCAACAGCAACGAGATAGT 360
DB 127 SerSerArgPro-----GlnAspArg 133
QY 361 CACGAGAGTGCACCGTTTCGATGAGGGTGATCTATTCCAGTTCACCGGTGTGCT 420
DB 134 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 153
QY 421 TTCTGGCTTACCAACACCAACGACACTGATGTTGTTGCTGTTCTCTACTGACACCAAC 480
DB 154 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 173
QY 481 AACACGACCAACGAGTGTGATCAAGTCCCGAGAGATTCATTTGGCTGGGAACACGGAG 540
DB 174 SerPheGlnAsnGlnLeuAspGlnMetProArgPheTyrLeuAlaGlyAsnGlnGlu 193
QY 541 CRAAGTTCTTAAGGTACCAACCAACAAAGCAGACAAAGCAGACGAGACGACGATACCAT 600
DB 194 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGly----- 206
QY 601 AGCCCATACAGCCCGCAAAAGTACGCTAGACGAGAGCGTGAAATTTAGCCCTCGAGGA 660
DB 207 -----Gly 207
QY 661 CAGCACAGCCGACGAGAACGAGCAGGACGAGACAGACAGACGAGTGAACATCTTC 720
DB 208 ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluGlnGlyGlySerIleLeu 227
QY 721 AGCGCTTACGCGGAGTTCTCGACAAAGCCCTCCAGGTTGACGACACAGATAGTG 780
DB 228 SerGlyPheAlaProGluPheLeuGluHisAlaPheVal---AspArgGlnIleVal 246
QY 781 CAAAACTTAAGAGCGCAGACCGAGAGTGAAGAGAGGGCCATTGTGACAGTGAAGGGA 840
DB 247 ArgLysLeuGlnGlyGluAsnGluGluGluGluGlyAlaIleValThrValLysGly 266
QY 841 GGCTCTAGATCTTTGACCCA-----GATAGAAAGAGACGTGCCGACGAGAAAGAG 891
DB 267 GlyLeuSerValIleSerProThrGluGluGlnGlnArgProGluGluGluGlu 286
QY 892 GAATACGATGAAGATGAATGAATACGATGAAGAGGATAGAAGCGGTGCGAGGGAAGC 951
```

```
DB 287 LysProAsp-----CysaspGluLysAspLysHisCysGlnSerGlnSer 301
QY 952 AGAGGCAGGGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTCTTAAAGAACATTT 1011
DB 302 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 318
QY 1012 GGTAGAAACAGATCCCTCGACATCTACAAACCTCAAGCTGGTTCCTCAAAATGCCAAC 1071
DB 319 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr 338
QY 1072 GATCTCAACCTTCTAATACTTAGTGGCTTGGACCTAGTCTGCAATATATGAAATCTCTAC 1131
DB 339 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 358
QY 1132 AGGAATGCAATGTTTGTGCTCCTACACACCAACGACACAGCATCATATATCGATTG 1191
DB 359 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 378
QY 1192 AGGGACGGGCTCAGTGCAGTCTGGACAGCAACGCGCAACAGAGTGTACGACGAGGAG 1251
DB 379 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 398
QY 1252 CTTCAAGAGGGTCACGTGCTTGTGTGTCACAGAACTTCGCGCTCGCTGAAAGTCCCCAG 1311
DB 399 LeuGlnGluGlyValValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 418
QY 1312 AGCGAGAACTTCAATACGTGCATTCAGACAGACAGCTCAAGGCCACGATAGCAACCTC 1371
DB 419 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 438
QY 1372 GCGGTGAAACTCCGTATAGATACCTGCGGAGGAGGTGGTTCAAATTCATATGCG 1431
DB 439 AlaGlyAlaAsnSerLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 458
QY 1432 CTCCTCAAGGAGCAGGCAAGGCGCTTAAGAACCAACCCCTTCAAGTTCTTCGTGCCA 1491
DB 459 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 478
QY 1492 CCGTCTCAGACAGTCTCCGAGGGCTGTGGCT 1521
DB 479 ProLysGluSerGlnArgValValAla 488
RESULT 11
US-10-425-114-44047
; Sequence 44047, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 44047
; LENGTH: 491
; TYPE: PR1
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700763253_FLI.pep
US-10-425-114-44047
Alignment Scores: 2,14e-120 Length: 491
Pred. No.: 1502.00 Matches: 297
Percent Similarity: 71.6% Conservatives: 68
Best Local Similarity: 58.2% Mismatches: 91
```


Db 34 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 53
QY 61 AATCGATTGAATCAGAGGCGGTATCATGTAGACTTGGAAACCCCAACACAGGAGTTC 120
Db 54 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 73
QY 121 GAATGCGCGCGCTCCCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGCCCT 180
Db 74 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 93
QY 181 TTCTACTCAATCTCCCAAGAGATCTTCATCCAGCAAGGAAGGGGATACTTTGGTTG 240
Db 94 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 113
QY 241 ATATTCCTCGTTGCTTAGACACTATGAAGACGCTCACACAAGGTGCTGCATCTCAG 300
Db 114 IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln 130
QY 301 TCCCAAGACCAACCAAGACGCTCTCCAAGGAGAGACCAAGCCAAACAGCAACGAGATAGT 360
Db 131 SerSerArgPro-----GlnAspArg 137
QY 361 CACCAAGAGTGCACCGTTTCGATGAGGGTGATCTCATTCAGTTCACACCGGTGTGCT 420
Db 138 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 157
QY 421 TTCCTGCTTACACGACCAACGACACTGATGTTGTGCTGTTCTTCTACTGACACCAAC 480
Db 158 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 177
QY 481 AACCAACGACACGCTTGATGATGTTCCCGAGGAGATTCATTTGGCTGGACACGGAG 540
Db 178 SerPheGlnAsnGlnLeuAspGlnMetProArgPheTyrLeuAlaGlyAsnGlnGlu 197
QY 541 CAAGAGTTCCTTAAGGTACCAACCAAGCAACAAAGCAGACAGAGAGAGCTTTACCATAT 600
Db 198 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGly----- 210
QY 601 AGCCCATACAGCCCGCAAGTCAGCTCAGCAAGAGAGAGGCTGAATTAGCCCTCAGGA 660
Db 211 -----Gly 211
QY 661 CAGCACAGCGCAGAGAACGAGCAGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 212 ThrGlnSerGlnLysGlyLysArgGlnGlnGlnGluGlnGlnGlyGlySerIleLeu 231
QY 721 AGCGGCTTACGCGGAGTTCCTGGAAACAAGCCTTCAGGTTGACGACACACAGATAGT 780
Db 232 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal--AspArgGlnIleVal 250
QY 781 CAAACCTAAGAGCGGACCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 251 ArgLysLeuGlnGlyGluAsnGluGluGluGluGlyGlyAlaIleValThrValLysGly 270
QY 841 GGCCTCAGATCTTGAGCCCA-----GATGAGAGAGAGTGCAGCAGAGAGAG 891
Db 271 GlyLeuSerValIleSerProThrGluGluGlnGlnGlnArgProGluGluGluGlu 290
QY 892 GAATACGATGAAGATGAATATGAATACGATCAAGAGGATAGAAGGCGTGCAGGGGAAGC 951
Db 291 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 305
QY 952 AGAGCGAGGGGGAATGTTATGAAGAGACATCTGCACCCCAAGTGTAAAAAGACATT 1011
Db 306 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 322
QY 1012 GGTAGAAACAGATCCCTCGACATCTACACCCCTCAAGCTGTTCACTCAAACTGCCAAC 1071
Db 323 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr 342
QY 1072 GATCTCAACCTCTTAATCTAGGTGGCTTGGACCTAGTCTGAATATGAAATCTCTAC 1131
Db 343 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 362

RESULT 13

US-10-425-114-46294
; Sequence 46294, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46294
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700757865_FLI.pep
US-10-425-114-46294

Alignment Scores:

Pred. No.:	2,14e-120	Length:	492
Score:	1502.00	Matches:	297
Percent Similarity:	71.6%	Conservative:	68
Best Local Similarity:	58.2%	Mismatches:	91
Query Match:	54.0%	Indels:	54
DB:	4	Gaps:	7

US-10-728-323-3 (1-1524) x US-10-425-114-46294 (1-492)

QY 1 CGSAGCAACCGGAGGAGAACGCGTCCAGTTCACGCGCTCAATCGCAGAGACCTGAC 60
Db 34 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp 53
QY 61 AATCGATTGAATCAGAGGCGGTTCATTTAGACTTGGAAACCCCAACACAGGAGTTC 120
Db 54 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 73
QY 121 GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGCCCT 180


```
QY 361 CACCAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCAGTTCACCAACCGGTGTGCT 420
Db 108 HisGlnLysValHisArgPheArgGluGlyAspLeuIleAlaValProThrGlyValAla 127
QY 421 TTCTGGCTCTACACGACACGACACTGATGTTGTTGTTCTTCTTACTGACACCAAC 480
Db 128 TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleIleAspThrAsn 147
QY 481 AACACGACAAACAGCTTCATCAGTTCCTCCAGAGATTCAATTTCGCTGGGAACACGAG 540
Db 148 SerLeuGluAsnGlnLeuAspGlnMetProArgPheTyrLeuAlaGlyAsnGlnGlu 167
QY 541 CAAGAGTTCCTTAAGGTACCAGCAACAAAGCAGACAAAGCAGACGAAAGCTTTACCATAT 600
Db 168 GlnGluPheLeuLysTyrGlnGln-----GlnGlnGlyGlySerGlnLysGly 176
QY 601 AGCCCATACAGCCCGCAAAAGTCAGCTAGACAAAGAGCGGTGAATTAGCCCTCAGGA 660
Db 177 -----GlnGlnGlyGlySerGlnLysGly 186
QY 661 CAGCAGCGCCGAGAGACGAGCAGGACAGAGAGAGAGAAACGAGGTGGAACATCTTC 720
Db 187 LysGln-----GlnGluGluGluAsnGluGlySerAsnIleLeu 199
QY 721 AGCGGCTTCACGCGGAGTTCCTGGAAACAGCTTCACGCTTGACACACAGACAGATAGTG 780
Db 200 SerGlyPheAlaProGluPheLeuLysGluAlaPheGlyVal---AsnMetGlnIleVal 218
QY 781 CAAACCTTAAGAGCGGAGACCGAGAGTGAAGAGAGGAGGCCATTGTGACGTAGGGGA 840
Db 219 ArgAsnLeuGlnGlyGluAsnGluGluAspSerGlyAlaIleValThrValLysGly 238
QY 841 GGCCTCAGATCTTGAGCCCGACAGATAGAAAGAGCGTGCAGACGAGAGAGGAATACGAT 900
Db 239 GlyLeuArgValThrAlaProAlaMetArgLysProGlnGlnGluAspAspAsp 258
QY 901 GAAGATGAA-----TATGAATACGATGAAGAGGATAGAGCGGTGGCAGGGGA 948
Db 259 GluGluGluGlnProGlnCysValGluThrAspLysGlyCysGlnArgGlnSerLysArg 278
QY 949 AGCAGAGCGGCGGGAATGGTATTGAAGAGAGATCTGCCCGCAAGCTCTAAAGAAC 1008
Db 279 SerArg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgGlnAsn 295
QY 1009 ATTGGTAGAAGACAGATCCCTCAGATCTACACCTCAGCTGTTCTACTCAAACTGCC 1068
Db 296 IleGlyGlnAsnSerSerProAspIleTyrAsnProGlnAlaGlySerIleThrAla 315
QY 1069 AACGATCTCAACCTTCTAATATCTTAGGTGGCTTGGACCTAGTGCTGAATATGAAATCTC 1128
Db 316 ThrSerLeuAspPheProAlaLeuTrpLeuLeuLysLeuSerAlaGlnTyrGlySerLeu 335
QY 1129 TACAGGAATGCAATTGTTGCTACTACACACCAACGACACAGCATCATATATCGA 1188
Db 336 ArgLysAsnAlaMetPheValProHisTyrThrLeuAsnAlaAsnSerIleIleTyrAla 355
QY 1189 TTGAGGGGACGGCTCACGTGCAAGTGGAGCAGACGACGCAACAGAGTGTACGACGAG 1248
Db 356 LeuAsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGly 375
QY 1249 GAGCTTCAAGAGGGTCACGTGTTGTGGTGCACAGAACTTCGCGCTCGCTGGAAAGTCC 1308
Db 376 GluLeuGlnGluGlyGlyValLeuIleValProGlnAsnPheAlaValAlaLysSer 395
QY 1309 CAGAGCGGAATTCGAATACGTGGCATTCAAGACAGACTCAAGGCCAGCATAGCCAAC 1368
Db 396 GlnSerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsn 415
QY 1369 CTCGCGGTGAAACTCCGTCATAGATAACCTGCGGAGGAGGTGGTTCGAATTCATAT 1428
Db 416 LeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnHisThrPhe 435
```

```
QY 1429 GGCCTTCAAAGGAGCAGGCAAGGCGCTTAAGAACAAACCCCTTCAAGTCTTCGTT 1488
Db 436 AsnLeuLysSerGlnGlnAlaArgGlnValLysAsnAsnProPheSerPheLeuVal 455
QY 1489 CCACGCTCTCAGCAGTCTCCGAGGGCTGTGGCT 1521
Db 456 ProProGlnGluSerGlnArgArgAlaValAla 466
```

Search completed: May 16, 2006, 00:00:55
Job time : 240.065 secs

THIS PAGE LEFT BLANK

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 23:41:32 ; Search time 5.24287 Seconds
(without alignments)
4094.130 Million cell updates/sec

Title: US-10-728-323-3

Perfect score: 2779

Sequence: 1 cggcagcaacggaggagaa.....ctccagggtgtggcttaa 1524

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 250354 seqs, 46948837 residues

Total number of hits satisfying chosen parameters: 500708

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlpx
-MODEL=frame+ n2p model
-Q=/abss/ABSSWEB_spool/US10728323/runat_15052006_172148_22650/app_query.fasta.1
-DB=Published Applications AA New -QFMT=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs802p
-USER=US10728323@CGN_1_1_63 @runat_15052006_172148_22650 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA New.*

1: /SIDSS/ptodata/1/pubpaa/US08_NEW_PUB.pepl.*
2: /SIDSS/ptodata/1/pubpaa/US06_NEW_PUB.pepl.*
3: /SIDSS/ptodata/1/pubpaa/US07_NEW_PUB.pepl.*
4: /SIDSS/ptodata/1/pubpaa/US08_NEW_PUB.pepl.*
5: /SIDSS/ptodata/1/pubpaa/PCT_NEW_PUB.pepl.*
6: /SIDSS/ptodata/1/pubpaa/US09_NEW_PUB.pepl.*
7: /SIDSS/ptodata/1/pubpaa/US09_NEW_PUB.pepl.*
8: /SIDSS/ptodata/1/pubpaa/US10_NEW_PUB.pepl.*
9: /SIDSS/ptodata/1/pubpaa/US10_NEW_PUB.pepl.*
10: /SIDSS/ptodata/1/pubpaa/US11_NEW_PUB.pepl.*
11: /SIDSS/ptodata/1/pubpaa/US11_NEW_PUB.pepl.*
12: /SIDSS/ptodata/1/pubpaa/US60_NEW_PUB.pepl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2673	96.2	507	11	US-11-033-039-96
2	1502	54.0	481	9	US-10-508-263-28
3	1499.5	54.0	485	9	US-10-508-263-22
4	1496	53.8	495	9	US-10-508-263-20
5	973.5	35.0	562	9	US-10-508-263-24
					Sequence 96, Appl
					Sequence 28, Appl
					Sequence 22, Appl
					Sequence 20, Appl
					Sequence 24, Appl

6	926.5	33.3	516	9	US-10-508-263-26	Sequence 26, Appl
7	779.5	28.0	472	9	US-10-508-263-62	Sequence 62, Appl
8	776.5	27.9	470	9	US-10-508-263-64	Sequence 64, Appl
9	773	27.8	493	9	US-10-508-263-30	Sequence 30, Appl
10	739.5	26.6	490	9	US-10-508-263-10	Sequence 10, Appl
11	739.5	26.6	509	9	US-10-508-263-18	Sequence 18, Appl
12	734.5	26.4	455	9	US-10-508-263-66	Sequence 66, Appl
13	712.5	25.6	488	9	US-10-508-263-12	Sequence 12, Appl
14	696.5	25.1	496	9	US-10-508-263-14	Sequence 14, Appl
15	689.5	24.8	451	9	US-10-508-263-68	Sequence 68, Appl
16	689.5	24.8	451	9	US-10-508-263-70	Sequence 70, Appl
17	682.5	24.6	524	9	US-10-508-263-60	Sequence 60, Appl
18	351	12.6	184	9	US-10-508-263-16	Sequence 16, Appl
19	269.5	9.7	358	11	US-11-096-568A-22915	Sequence 22915, A
20	269.5	9.7	360	11	US-11-096-568A-22914	Sequence 22914, A
21	269.5	9.7	372	11	US-11-096-568A-22913	Sequence 22913, A
22	233	8.4	360	11	US-11-096-568A-25849	Sequence 25849, A
23	233	8.4	385	11	US-11-096-568A-25848	Sequence 25848, A
24	220	7.9	329	11	US-11-096-568A-17189	Sequence 17189, A
25	217.5	7.8	325	11	US-11-096-568A-25850	Sequence 25850, A
26	211	7.6	352	11	US-11-096-568A-12807	Sequence 12807, A
27	211	7.6	358	11	US-11-096-568A-12806	Sequence 12806, A
28	170.5	6.1	626	9	US-10-498-036-89	Sequence 89, Appl
29	170.5	6.1	626	11	US-11-033-039-10	Sequence 10, Appl
30	147.5	5.3	639	9	US-10-821-234-907	Sequence 907, Appl
31	140.5	5.1	422	11	US-11-096-568A-6597	Sequence 6597, Ap
32	140.5	5.1	458	11	US-11-096-568A-6596	Sequence 6596, Ap
33	140	5.0	901	11	US-11-087-099-1818	Sequence 1818, Ap
34	138.5	5.0	234	11	US-11-096-568A-12808	Sequence 12808, A
35	137.5	4.9	1229	11	US-11-188-298-7436	Sequence 7436, Ap
36	132.5	4.8	761	9	US-10-485-517-252	Sequence 252, App
37	130	4.7	480	11	US-11-096-568A-249	Sequence 249, App
38	130	4.7	480	11	US-11-096-568A-251	Sequence 251, App
39	130	4.7	520	11	US-11-096-568A-248	Sequence 248, App
40	123.5	4.7	715	11	US-11-072-512-3385	Sequence 3385, Ap
41	129	4.6	1609	11	US-11-087-099-4181	Sequence 4181, Ap
42	128	4.6	1971	9	US-10-961-231-1	Sequence 1, Appli
43	128	4.6	1971	11	US-11-179-624-1	Sequence 1, Appli
44	124.5	4.5	625	9	US-10-784-094-387	Sequence 387, App
45	123	4.4	313	11	US-11-096-568A-16091	Sequence 16091, A

ALIGNMENTS

RESULT 1
US-11-033-039-96
; Sequence 96, Application US/11033039
; Publication No. US20060002947A1
; GENERAL INFORMATION:
; APPLICANT: HUMPHREYS, ROBERT
; APPLICANT: XU MINZHEN
; TITLE OF INVENTION: LI-KEY/ANTIGENIC EPIOTOPE HYBRID PEPTIDE VACCINES
; FILE REFERENCE: REH-2017US01
; CURRENT APPLICATION NUMBER: US/11/033,039
; CURRENT FILING DATE: 2005-01-11
; PRIOR APPLICATION NUMBER: 10/245,871
; PRIOR FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: 10/197,000
; PRIOR FILING DATE: 2002-07-17
; PRIOR APPLICATION NUMBER: 09/396,813
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 1452
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 96
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Arachis hypogaea
US-11-033-039-96
Alignment Scores:
Pred. No.: 2.72e-220 Length: 507
Score: 2673.00 Matches: 507
Percent Similarity: 100.0% Conservative: 0

Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	96.2%	Indels:	0
DB:	11	Gaps:	0
US-10-728-323-3 (1-1524) x US-11-033-039-96 (1-507)			
QY	1	CGGACCAACCGGAGAGAACCGTCCAGTTCCAGCGCTCAATGCGCAGAGACTGAC	60
DB	1	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	20
QY	61	AATCGATTGAATCAGAGCGCGGTACATTGAGACTTGGAAACCCCAACACAGAGAGTTC	120
DB	21	AsnArgLleGluSerGluGlyGlyTyrlleGluThrTrpAsnProAsnAsnGlnGluPhe	40
QY	121	GAATGCGCGCGGTCCCTCTCTCGTTAGTCTCCGCGCAACGCCCTTCGTAGCGCT	180
DB	41	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro	60
QY	181	TTCTACTCAATGCTCCCGAGGATCTTCATCCAGCAAGGAGGGGATCTTGGGTTG	240
DB	61	PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu	80
QY	241	ATATTCCCTGGTTGCTTAGACACTATGAAGAGCCTCACACAAGGTGCTGATCTCAG	300
DB	81	LlePheProGlyCysProArgHisTyrGluGluProHisThrGlnGlyArgArgSerGln	100
QY	301	TCCAAAGACCAACAGAGCTCTCCAAGGAGAAGCAACAAAGCCACAGCAACGAGATGT	360
DB	101	SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnGlnArgAspSer	120
QY	361	CACCAGAGGTGACCGTTTCAGTAGGGTGATCTTCATTCAGTTCCACCGGTGTGCT	420
DB	121	HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	140
QY	421	TTCTGTGCTCTACAAACACACAGACTGATGTGTGTGCTGTCTCTTACTGACACCAAC	480
DB	141	PheTrpLeuTyrAsnAspHisAspThrAspValAlaValAsnLeuThrAspThrAsn	160
QY	481	AACAACGACCAACAGCTTGATCAGTTCCCGAGAGATTCAATTTGGCTGGCAACCGAG	540
DB	161	AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnThrGlu	180
QY	541	CAAGAGTTCTTAAGTACGACGACCAACAGCAAGCAAGCAGCAGAGAGCTTACCATAT	600
DB	181	GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgArgSerLeuProTyr	200
QY	601	AGCCCATACAGCCGCAAGTCCAGCTAGACAAGAGAGCGTGAATTTAGCCCTCAGAGA	660
DB	201	SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly	220
QY	661	CAGCACGCGCAGAGACGAGCAGCAAGCAAGAGAGAAACCAAGGTGGAACATCTTTC	720
DB	221	GlnHisSerArgArgGluArgAlaGlyGlnGluGluGluAsnGluGlyGlyAsnIlePhe	240
QY	721	AGCGGCTTACGCGGAGTCTCCGGAACAAGCCCTTCAGTTGACACACAGAGATAGC	780
DB	241	SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal	260
QY	781	CAAACTTAAGAGCGCAGACCGAGAGTGAAGAGAGGAGCCATTGTGACAGTGAAGGGA	840
DB	261	GlnAsnLeuArgGlyGluThrGluSerGluGluGluGlyAlaIleValThrValArgGly	280
QY	841	GGCCTCAGATCTTGAAGCCAGATAGAAAGAGACGTGCCGACCAAGAGAGGAAATACAT	900
DB	281	GlyLeuArgIleLeuSerProAspArgLysArgArgAlaAspGluGluGluTyrAsp	300
QY	901	GAGATGAATATCAATCAGTGAAGGATAGAGCGGTGGCAGGGGAGCCAGGACGAG	960
DB	301	GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg	320
QY	961	GGGAATGGTATTCAAGACAGCATCTGCACCGCAAGTGTAAAAAGAACATTGTTAGAAC	1020
DB	321	GlyAsnGlyLleGluGluThrIleCysThrAlaSerAlaLysLysAsnIleGlyArgAsn	340
QY	1021	AGATCCCCTGACATCTACAACCCCTCAAGCTGGTCTCACTCAAACTGCCAACGATCTCAAC	1080
DB	341	ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn	360
QY	1081	CTTCTAAACTTATGAGTGGCTTGGACCTAGTGTGTAATATGGAAATCTTACAGGAATGCA	1140
DB	361	LeuLeuIleLeuArgTrpLeuGlyProSerAlaGluTyrGlyAsnLeuTyrArgAsnAla	380
QY	1141	TTGTTTGTGCTCCTACAAACACCAACGACACAGCATCATATATCGATTGAGGGGACGG	1200
DB	381	LeuPheValAlaHisTyrAsnThrAsnAlaHisSerIleIleTyrArgLeuArgGlyArg	400
QY	1201	GCTCAGCTCAAGTCTGTGACAGCAACCGCAACAGAGTGTACGACGAGAGCTTCAAGG	1260
DB	401	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu	420
QY	1261	GGTCAAGTCTGTGTCGACAGAACTTCGCGCTCGCTGGAAAGTCCCAGAGCGGAGAAC	1320
DB	421	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAsn	440
QY	1321	TTTCAATACGTGGCATTTCAAGACAGATCAAGGCCGAGCATAGCCAACTTCGCGGTGAA	1380
DB	441	PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	460
QY	1381	AACTCGCTCATAGATAACTCGCGAGAGGTGTTGCAAAATTCATATGCGCTCCAAAG	1440
DB	461	AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg	480
QY	1441	GAGCAGGCAAGGCAGCTTAAGAACCAACACCCCTTCAAGTCTTCCGCTCCACGCTCAG	1500
DB	481	GluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProProSerGln	500
QY	1501	CAGTCTCCGAGGCTGTGCT 1521	
DB	501	GlnSerProArgAlaValAla 507	
RESULT 2			
US-10-508-263-28			
; Sequence 28, Application US/10508263			
; Publication No. US20050260754A1			
; GENERAL INFORMATION:			
; APPLICANT: BASF Plant Science GmbH			
; TITLE OF INVENTION: Constructs and methods for regulating gene expression			
; FILE REFERENCE: 53262-20085.00			
; CURRENT APPLICATION NUMBER: US/10/508,263			
; CURRENT FILING DATE: 2004-09-20			
; NUMBER OF SEQ ID NOS: 126			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 28			
; LENGTH: 481			
; TYPE: PRT			
; ORGANISM: Glycine max			
US-10-508-263-28			
Alignment Scores:			
Pred. No.:	3,37e-120	Length:	481
Score:	1502.00	Matches:	297
Percent Similarity:	71.6%	Conservative:	68
Best Local Similarity:	58.2%	Mismatches:	91
Query Match:	54.0%	Indels:	54
DB:	9	Gaps:	7
US-10-728-323-3 (1-1524) x US-10-508-263-28 (1-481)			
QY	1	CGGACCAACCGGAGAGAACCGTCCAGTTCCAGCGCTCAATGCGCAGAGACTGAC	60
DB	23	ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnArgLeuAsnAlaLeuLysProAsp	42
QY	61	AATCGATTGAATCAGAGCGCGGTATCATTTGAGACTTGGAAACCCCAACACAGGAGTTC	120
DB	43	AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe	62

```
QY 121 GAATGCGCGCGTCCGCTCTCTCGCTTAGTCTCCGCGCCGCAACGCGCTTCGTAGGCT 180
Db ::::::::::::::::::::
63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 82
QY 181 TTCTACTCAAGTCCTCCAGGAGATCTTCATCCAGCAGGAGGAGGATCTTCGGTTG 240
Db ::::::::::::::::::::
83 SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 102
QY 241 ATATTCCCTCGTGTCTAGACACTATGAAGAGCCTTCACACAGGCTCGTCTGATCTCAG 300
Db ::::::::::::::::::::
103 IlePheProGlyCysProSerThrPheGluGluProGlnGlnGlySerGly-----Gln 119
QY 301 TCCCAAGACACCAAGAGCTCTCAAGGAGAGAACCAAGCAACAGCAAGCAGATAGT 360
Db ::::::::::::::::::::
120 SerSerArgPro-----GlnAspArg 126
QY 361 CACCAAGAGGTGCAACGGTTTCGATGAGGTGTATCTCATTCAGTTCACCGGTGTTGCT 420
Db ::::::::::::::::::::
127 HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146
QY 421 TTCTGGCTCTACAACGACACCACTGATGTGTGTGCTTCTCTTCTTCTTCTTCTTCTTCT 480
Db ::::::::::::::::::::
147 TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166
QY 481 AACAAACGACCAACAGCTTGATCAGTTCCTCCAGGAGATTCATTTCCGCTGGGAACAGGAG 540
Db ::::::::::::::::::::
167 SerPheGlnAsnGlnLeuAspGlnMetProArgPheTyrLeuAlaGlyAsnGlnGlu 186
QY 541 CAAGAGTCTTAAAGTACCAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT 600
Db ::::::::::::::::::::
187 GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly----- 199
QY 601 AGCCCATACAGCCCGCAAGTCAGCTAGACAAAGAGCGTGAATTTAGCCCTCGAGGA 660
Db ::::::::::::::::::::
200 -----Gly 200
QY 661 CAGCACAGCCGACAGAGACGAGCAGGACAAAGAAAGAAAGAAAGAAAGTTC 720
Db ::::::::::::::::::::
201 ThrGlnSerGlnLysGlyLysArgGlnGlnGluGlnGluGlnGluGlnGlySerIleLeu 220
QY 721 AGCGGCTTCACCGCGGATCTCTGAAACAAAGCTTCAGGTTCCAGGTTGACGACGAGATGTG 780
Db ::::::::::::::::::::
221 SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239
QY 781 CAAACCTTAAGAGCGGACGACGAGTGAAGAGAGGAGGAGGATTCGACAGTGAAGGGA 840
Db ::::::::::::::::::::
240 ArgLysLeuGlnGlyLysGlnGluGluGluGlyAlaIleValThrValLysGly 259
QY 841 GCGCTCAGAAATCTTGAGCCCA-----GATAGAAAGAGACGTGCGCGACGAAGAAGAG 891
Db ::::::::::::::::::::
260 GlyLeuSerValIleSerProProThrGluGluGlnGlnGlnArgProGluGluGluGlu 279
QY 892 GAATACGATGAAGATGAATATGAATACGATGAAGAGATAGAAAGCGGTGGCAGGGGAAGC 951
Db ::::::::::::::::::::
280 LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294
QY 952 AGAGCAGGGGGAATGTTTGAAGAGAGATCTCCAGCGCAAGTCTCAATAAAGACATTT 1011
Db ::::::::::::::::::::
295 Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311
QY 1012 GGTAGAAACAGATCCCTTCACATCTACAAACCTCAAGCTGGTTCCTCAAAACTCCCAAC 1071
Db ::::::::::::::::::::
312 GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr 331
QY 1072 GATCTCAACCTTCTAATATAGTGGCTTGACCTAGTGTGAAATGGAATCTCTAC 1131
Db ::::::::::::::::::::
332 SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg 351
QY 1132 AGGAATGATCTTGTTCCTCCTACACACCAAGCAGCAGCATATATATCGATTG 1191
Db ::::::::::::::::::::
352 LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371
QY 1192 AGGGGACGGGCTCACGTGCAAGTCTGGACAGACCAACGCGCAACAGAGTGTACGACGAGGAG 1251
```

```
Db 372 AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 391
QY 1252 CTTCAAGAGGCTCAGCTGCTTGTGTGTCACAGAACTTCGCGCTCGCTGGAAGTCCCGAG 1311
Db 392 LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411
QY 1312 AGCGAGAACTTCGAATACGTGCGCATTCAGACAGACAGCTCAAGGCCCGACATAGCCAACTTC 1371
Db 412 SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431
QY 1372 GCGGTGAAACTCGCTCATAGATAACCTGCGGAGGAGGTGGTTCGAAATTCATATGCG 1431
Db 432 AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 451
QY 1432 CTCCAAAGGGAGCAGGCAAGCAGCTTAAGAAACAAACACCCCTTCACAGTTCCTTCCTCCA 1491
Db 452 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 471
QY 1492 CCGTCTCAGCAGCTCTCCGAGGCTGTGGCT 1521
Db 472 ProLysGluSerGlnArgArgValValAla 481

RESULT 3
US-10-508-263-22
; Sequence 22, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Glycine max
US-10-508-263-22

Alignment Scores:
Pred. No.: 5,53e-120 Length: 485
Score: 1499.50 Matches: 297
Percent Similarity: 72.0% Conservative: 71
Best Local Similarity: 58.1% Mismatches: 94
Query Match: 54.0% Indels: 49
DB: 9 Gaps: 6

US-10-728-323-3 (1-1524) x US-10-508-263-22 (1-485)
QY 1 CGGCAGCAACCGGAGGAGAACGCGTCCAGTTCACGCGCTCAATGCCAGACGACTGAC 60
Db 20 ArgGluGlnAlaGlnGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp 39
QY 61 AATCGCATTAATCAGAGGCGGTTTACATTTCAGACTTCGAAACCCCAACAAACAGGAGTTC 120
Db 40 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 59
QY 121 GAATGCGCGCGCTGCGCTCTCTCGCTTAGTCTCTCCGCGCAACGCGCTTCGTAGGCT 180
Db 60 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 79
QY 181 TTCTACTCAATGCTCCCGAGGATCTTCATCCAGCAAGGAGGAGGATCTTCGGTTG 240
Db 80 SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyAsnGlyIlePheGlyMet 99
QY 241 ATATTCCCTCGTGTCTAGACACTATGAAGAGCCTTCACACAGGCTCAGACAGGTCGTGCTCAG 300
Db 100 IlePheProGlyCysProSerThrTyrGlnGluProGlnGlnGluSerGlnGlnArgGlyArg 119
QY 301 TCCCAAGACCAACCAAGACGCTCTCCAGGAGAGAACCAAGCAACAGAGTGTACGACGAGT 360
```

Db 120 SerGlnArgPro-----GlnAspArg 126
QY 361 CACCAGAGGTGACCGTTTCGATGAGGGTGATCTCATTGCGAGTTCGCCACCGGTGTGCT 420
Db 127 HisGlnLysValHisArgPheArgGluGlyAspLeuIleAlaValProThrGlyValAla 146
QY 421 TTCTGGCTCTAACGACACACGACACTGATGTTGTGCTGTTCTTCTTACTGACACCAAC 480
Db 147 TrpTrpMetTyrAsnAsnGluAspThrProValAlaValSerIleIleAspThrAsn 166
QY 481 AACACGCAACACAGCTTCATCAGTTCCCGAGGAGATTCAATTTGGCTGGGACACGGAG 540
Db 167 SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
QY 541 CAAAGAGTTCTTAAGGTACAGCAACAAAGCAGACAAAGCAGACGAGAAAGCTTACCATAT 600
Db 187 GlnGluPheLeuLysTyrGlnGln----- 195
QY 601 AGCCCATACAGCCCGCAAGTCAAGCTAGCAGCAAGAGAGCGTGAAATTTAGCCCTCGAGGA 660
Db 196 -----GlnGlnGlySerGlnSerGlnLysGly 205
QY 661 CAGCACAGCCGACAGAACGACGACGACAAAGAAAGAAAGAAAGAGGTGGAAACATCTTC 720
Db 206 LysGln-----GlnGluGluAsnGluGlySerHisIleLeu 218
QY 721 AGCGGTTCCAGCCGAGTTCCTGGAAACAGCCTTCAGGTTGACGACGACAGATAGTG 780
Db 219 SerGlyPheAlaProGluPheLeuLysGluAlaPheGlyVal---AsnMetGlnIleVal 237
QY 781 CAAACCTTAAGGCGCAGACCGAGTGAAGAGAGGAGCCATTTGACAGTGGGGCA 840
Db 238 ArgAsnLeuGlnGlyGluAsnGluGluAspSerGlyAlaIleValThrValLysGly 257
QY 841 GGCTCAGAAATCTGAGCCAGATAGAAAGAGACGTCGCCGACGAAAGAGAGAAATACGAT 900
Db 258 GlyLeuArgValThrAlaProAlaMetArgLysProGlnGlnGluAspAspAsp 277
QY 901 GAAGATGAA-----TATGAATACGATGAAGAGGATAGAAGCGGTGGCAGGGGA 948
Db 278 GluGluGluGlnProGlnCysValGluThrAspLysGlyCysGlnArgGlnSerLysArg 297
QY 949 AGCAGAGGAGGGGAATGGTATTGAAGAGACGATCTGCACCCAGAGTCTAAAAAGAAC 1008
Db 298 SerArg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgGlnAsn 314
QY 1009 ATTGGTAGAAACAGATCCCTCGACATCTCAACCTCAAGCTGGTTCACTCAAAAATGCC 1068
Db 315 IleGlyGlnAsnSerSerProAspIleTyrAsnProGlnAlaGlySerIleThrAla 334
QY 1069 AAGCATCTCAACCTTCTAATACATTAGGTGGCTTGGACCTAGTCTGAATATGGAATCTC 1128
Db 335 ThrSerLeuAspPheProAlaLeuTrpLeuLeuLysLeuSerAlaGlnTyrGlySerLeu 354
QY 1129 TACAGGAATGCATTGTTGCTGCTACTCAACACCAACGACACAGCATCATATATCGA 1188
Db 355 ArgLysAsnAlaMetPheValProHisTyrThrLeuAsnAlaAsnSerIleIleTyrAla 374
QY 1189 TTCAGGGGACGGCTCAGCTGCAAGTCGTGGACGACCAACGGCAACAGAGTGTACGACGAG 1248
Db 375 LeuAsnGlyArgAlaLeuValGlnValValAsnCysAsnGlnGluArgValPheAspGly 394
QY 1249 GAGCTTCAAGAGGTCACGCTTGTGTGGTCCACAGAACTTCGCCGTGCGTGGAAAGTCC 1308
Db 395 GluLeuGlnGluGlyValLeuIleValProGlnAsnPheAlaValAlaLaLysSer 414
QY 1309 CAGAGCGAGAACCTCGAATACGTGGCATTCACAGCAGACTCAAGGCCCGCATAGCCAAC 1368
Db 415 GlnSerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsn 434
QY 1369 CTCGCCGTGAAACCTCCGTATAGATAACCTCCGAGAGAGGTGGTTGCAAAATTCATAT 1428
Db 435 LeuAlaGlyAlaAsnSerLeuAsnAlaLeuProGluGluValIleGlnHisThrPhe 454

QY 1429 GGCCTCCAAAGGAGCAGGACGCTTAAGAACAAACCCCTTCAAGTTCCTCGTT 1488
Db 455 AsnLeuLysSerGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuVal 474
QY 1489 CCACCGTCTCAGCAGCTCTCCGAGGCGCTGGCT 1521
Db 475 ProProGlnGluSerGlnArgArgAlaValAla 485
RESULT 4
US-10-508-263-20
; Sequence 20, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Glycine max
US-10-508-263-20
Alignment Scores:
Pred. No.: 1,116-119 Length: 495
Score: 1496.00 Matches: 292
Percent Similarity: 70.9% Conservative: 73
Best Local Similarity: 56.7% Mismatches: 100
Query Match: 53.8% Indels: 50
DB: Gaps: 6
US-10-728-323-3 (1-1524) x US-10-508-263-20 (1-495)
QY 1 CGCAGCAACCGGAGAGAACCGTCCAGTTCAGCGCCTCAATGCCGAGACCTGAC 60
Db 23 ArgGluGlnProGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp 42
QY 61 AATCGCATTGAATCAGAGGCGGTTCATATTGAGACTTGGAAACCCCAACACGAGGTTC 120
Db 43 AsnArgIleGluSerGluGlyGlyLeuIleGluThrTrpAsnProAsnAsnLysProPhe 62
QY 121 GAATGCGCGCGCTCGCCTCTCTCGTTAGTCTCCCGCGCAACGCCCTTCGTAGGCCT 180
Db 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 82
QY 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGATACTTTGGTTG 240
Db 83 SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyLysGlyIlePheGlyMet 102
QY 241 ATATTCCCTGGTTGTCTAGACACTATGAAGAGCCTTCACACAAAGGTGCTCATCTCAG 300
Db 103 IleTyrProGlyCysProSerThrPheGluGluProGlnGlnProGlnArgGlyGln 122
QY 301 TCCCAAAAGACCAACAGCGTCTCCAAGGAGAAAGCAAAAGCCAAACGACGAGATAGT 360
Db 123 SerSerArgPro-----GlnAspArg 129
QY 361 CACCAGAGGTGCACCGTTTCGATGAGGGTGATCTCATTGCGAGTTCACCGGTGTGCT 420
Db 130 HisGlnLysIleTyrAsnPheArgGluGlyAspLeuIleAlaValProThrGlyValAla 149
QY 421 TTCTGGCTCTACAAGCAGCACACACACTGATGTTGTGTGCTTCTTCTTACTGACACCAAC 480
Db 150 TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleIleAspThrAsn 169
QY 481 AACACGCAACACAGCTTCATCAGTTCCCGAGGAGATTCAATTTGGCTGGGACACGGAG 540
Db 170 SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 189


```
Db 237 PheLeuAlaGlnSerPheAsnThrAsnGlu---AspIleAlaGluLysLeu-----Glu 253
QY ACCGAGAGTGAAGAAGAGGAGCCATTGTGCACAGTGGAGGGAGGCTCAGAAATCTTGAGC 858
  :: ::::
Db 254 SerProAspGluArgLysGlnIleValThrValGluGlyGlyLeuSerValIleSer 273
  :::::
QY 859 CCAGATAGAAGAGAGCTGCCGACGAGAGGAAGGAATACGATGAA----- 903
  :::::
Db 274 ProLysTrpGlnGlnGlnAspGluAspGluAspGluAspGluAspGluAspGlu 293
  :::::
QY 904 -----GATGAATAT 912
  :::::
Db 294 GlnIleProSerHisProProArgArgProSerHisGlyLysArgGluGlnAspGluAsp 313
  :::::
QY 913 GAATACGATGAAGAGATGAG---AGCGTGGCAGGGGAGAGCAGAGC----- 957
  :::::
Db 314 GluAspGluAspGluAspLysProArgProSerArgProSerGlnGlyLysArgAsnLys 333
  :::::
QY 957 ----- 957
Db 334 ThrGlyGlnAspGluAspGluAspGluAspGluAspGlnProArgLysSerArgGluTrp 353
  :::::
QY 958 -----AGGGGG--- 963
  :::::
Db 354 ArgSerLysLysThrGlnProArgArgProArgGlnGluProArgGluArgGlyCys 373
  :::::
QY 964 -----AATGGTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAGAACATGTGT 1014
  :::::
Db 374 GluThrArgAsnGlyValGluGluAsnIleCysThrLeuLysLeuHisGluAsnIleAla 393
  :::::
QY 1015 AGAAACAGATCCCTCACATCTACAACTTCAAGCTGGTCTCACTCAAACTGCCACGAT 1074
  :::::
Db 394 ArgProSerArgAlaAspPheTyrAsnProLysAlaGlyArgIleSerThrLeuAsnSer 413
  :::::
QY 1075 CTCAACTTCTAATACTTAGTGGCTTGGACCTAGTCTCAATATGCAATCTACAGG 1134
  :::::
Db 414 LeuThrLeuProAlaLeuArgGlnPheGlnLeuSerAlaGlnTyrValValLeuTyrLys 433
  :::::
QY 1135 AATGCATTTGTTCTCCTCCTACACACACGACAGCATCATATATATATCATGATTGAGG 1194
  :::::
Db 434 AsnGlyIleTyrSerProHisTrpAsnLeuAsnAlaAsnSerValIleTyrValThrArg 453
  :::::
QY 1195 GACGGGCTCACGTCAAGTCGTGGACAGCAACGCGCAACAGAGTGTACGACGAGGAGCTT 1254
  :::::
Db 454 GlyGlnGlyLysValArgValValAsnCysGlnGlyAsnAlaValPheAspGlyGluLeu 473
  :::::
QY 1255 CAAGAGGGTCACGTGCTTGTGTGCCACAGAACTTCGCGTCTCGTGGAAAGTCCACAGAGC 1314
  :::::
Db 474 ArgArgGlyGlnLeuLeuValValProGlnAsnPheValValAlaGluGlnAlaGlyGlu 493
  :::::
QY 1315 GAGAACTTGAATACGTGGCATTCAGACAGACTCAAGGCCAGCATAGCCACCTCGCC 1374
  :::::
Db 494 GlnGlyPheGluTyrIleValPheLysThrHisAsnAlaValThrSerTyrLeu--- 512
  :::::
QY 1375 GGTGAAACTCCTCGTCATAGTAACCTCCGGAGAGGTGGTGTGCAATTCATATGSCCTC 1434
  :::::
Db 513 -----LysAspValPheArgAlaIleProSerGluValLeuAlaHisSerTyrAsnLeu 530
  :::::
QY 1435 CAAAGGAGCAGCAAGGAGCTTAAAGAACAAACCCCTTCAAGTCTCTCGTT---CCA 1491
  :::::
Db 531 ArgGlnSerGlnValSerGluLeuLysTyrGluGlyAsnTrpGlyProLeuValAsnPro 550
  :::::
QY 1492 CCGTCTCAGCAG---TCTCCGAGG 1512
  :::::
Db 551 GluSerGlnGlnGlySerProArg 558
  :::::
```

RESULT 6

```
US-10-508-263-26
; Sequence 26, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
```

```
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Glycine max
US-10-508-263-26

Alignment Scores:
Pred. No.: 5,42e-71 Length: 516
Score: 926.50 Matches: 207
Percent Similarity: 53.1% Conservative: 72
Best Local Similarity: 39.4% Mismatches: 163
Query Match: 33.3% Indels: 83
DB: 9 Gaps: 11
```

```
US-10-728-323-3 (1-1524) x US-10-508-263-26 (1-516)
QY 19 AACGGCTGCCAGTTCCAGCGCTCAATGCGCAGAGACCTGCAATCGCAATCGCAATGAATCAGAG 78
  :::::
Db 31 AsnGluCysGlnLeuAsnAsnLeuAsnAlaLeuGluProAspHisArgValGluSerGlu 50
  :::::
QY 79 GCGGGTTACATTGAGACTTGGAAACCCCAACACACAGGAGTTCGAATGCCCGCGCTCGCC 138
  :::::
Db 51 GlyGlyLeuIleGluThrTrpAsnSerGlnHisProGluLeuGlnCysAlaGlyValThr 70
  :::::
QY 139 CTCCTCTCGTTCAGTCTCCGCGCAACGCCCTTCGTAGGCTTCTTACTCTCAATGCTCC 198
  :::::
Db 71 ValSerLysArgThrLeuAsnArgAsnGlySerHisLeuProSerTyrLeuProTyrPro 90
  :::::
QY 199 CAGGAGATCTTCATCAGCAAGAGGGGATACTTTGGGTTGATATTCCTGGTTGCTCT 258
  :::::
Db 91 GlnMetIleIleValValGlnGlyLysGlyAlaIleGlyPheAlaPheProGlyCysPro 110
  :::::
QY 259 AGACACTATGAAGAGCCTCACACACAGGTCTCGATCTCAGTCCCAAGACACCAAGA 318
  :::::
Db 111 GluThrPheGluLysProGlnGlnGlnSerSerArg----- 122
  :::::
QY 319 CGTCTCCAAGAGGAAGACCAAAAGCAACAGACAGATAGTACCAGAAAGGTGACCCGT 378
  :::::
Db 123 -----ArgGlySerArgSerGlnGlnGlnLeuGlnAspSerHisGlnLysIleArgHis 140
  :::::
QY 379 TTCGATAGGGTGATCTCATTCAGATTCCACCGGTGTTGCTTTCTGGCTCTTACAAACAC 438
  :::::
Db 141 PheAsnGluGlyAspValLeuValIleProLeuGlyValProTyrTrpThrTyrAsnThr 160
  :::::
QY 439 CACGACACTGATGTTGTTGCTCTTCTTCTTACTGACACCAACAAACACACACACAGCTT 498
  :::::
Db 161 GlyAspGluProValValAlaIleSerProLeuAspThrSerAsnPheAsnAsnGlnLeu 180
  :::::
QY 499 GATCAGTTCCTCCAGGAGATTCAATTTGGCTGGGAACACGAGCAAGAGTTCTTAAGGTAC 558
  :::::
Db 181 AspGlnAsnProArgValPheTyrLeuAlaGlyAsnProAspIleGlu----- 196
  :::::
QY 559 CAGCAACAAAGCAGACAAAGCAGACGAGAGCTTACCATATAGCCCTACAGCCCGCGAA 618
  :::::
Db 197 -----HisProGluThrMetGln 202
  :::::
QY 619 AGTCAGCTTAGACAAAGAGCGTGAATTTAGCCCTCGAGGACAGCAGCCGCGAGAA 678
  :::::
Db 203 GlnGlnGlnGlnGlnLysSerHisGlyGlyArgLysGlnGlyGlnHis----- 218
  :::::
QY 679 CGAGCAGGACAAGAAGAAAGAAACATCTTCAGCGGGTTCACGCGCGAG 738
  :::::
Db 219 -----ArgGlnGlnGlnGluGluGlyGlySerValLeuSerGlyPheSerLysHis 235
  :::::
QY 739 TTCCTCGAAACAGCCCTTCAGGTTGACGACACAGATAGTGCAAAACCTTAAGCGCGAG 798
  :::::
Db 236 PheLeuAlaGlnSerPheAsnThrAsnGlu---AspThrAlaGluLysLeuArg----- 252
  :::::
```

```
QY 799 ACCGAGAGTGAAGAAGAGGAGCCATTGTGACAGTGAGGGAGGCGCTCAGAACTTTGAGC 858
Db   :|||
QY 253 SerProAspGluArgLysGlnIleValThrValGluGlyGlyLeuSerValIleSer 272
Db   :|||
QY 859 CCAGATAGAAGAGACGTCGCCACGAAGAGAGGAATAGCATGAA----- 903
Db   :|||
QY 273 ProLysTrpGlnGluGlnGluAspGluAspGluAspGluGluTyrGlyArgThr 292
Db   :|||
QY 904 -----GATGAATATGAATAC 918
Db   :|||
QY 293 ProSerTyrProProArgArgProSerHisGlyLysHisGluAspGluAspGluAsp 312
QY 919 GATGAAGAGGATAGA----- 933
Db   :|||
QY 313 GluGluGluAspGlnProArgProAspHisProProGlnArgProSerArgProGln 332
QY 934 -----AGCGTGGCAGGGGAAGCAGAGCAGCGGGGAATGTATTGAAGAGCATCTGC 987
Db   :|||
QY 333 GlnGluProArgGlyArgGlyCysGlnThrArg---AsnGlyValGluGluAsnIleCys 351
QY 988 ACCGCAAGTGCTAAAGAACAATCGTAGAAACAGATCCCTGACATCTACAAACCTCAA 1047
Db   :|||
QY 352 ThrMetLysLeuHisGluAsnIleAlaArgProSerArgAlaAspPheTyrAsnProLys 371
QY 1048 GCTGTTCACTCAAAACCTGCCAAGCATCTCAACCTTTCTAATACTTAGTGGCTTGAACCT 1107
Db   :|||
QY 372 AlaGlyArgIleSerThrLeuAsnSerLeuThrLeuProAlaLeuArgGlnPheGlyLeu 391
QY 1108 AGTGCTGAATATGGAATCTTACAGGAATGCAATGTTTGTTCGCTCCTACCAACCAAC 1167
Db   :|||
QY 392 SerAlaGlnTyrValValLeuTyrArgAsnGlyIleTyrSerProAspTrpAsnLeuAsn 411
QY 1168 GCACACAGCATATATATGATTGAGGAGCGGGCTCAGTGCACAGTCTGTGGACAGCAAC 1227
Db   :|||
QY 412 AlaAsnSerValThr---MetThrArgGlyLysGlyArgValArgValValAsnCysGln 430
QY 1228 GCACACAGAGTACACAGAGAGGTTCAGAGGGTTCAGTCTTGTGTGTCACAGAAC 1287
Db   :|||
QY 431 GlyAsnAlaValPheAspGlyGluLeuArgArgGlyGlnLeuValValProGlnAsn 450
QY 1288 TTCGCGCTGCTGGAAGTCCAGAGCGAGAACTTCGAATAGTGGCATTCGAAGACAGAC 1347
Db   :|||
QY 451 ProAlaValAlaGluGlnGlyGlyGluGlnGlyLeuGluTyrValValPheLysThrHis 470
QY 1348 TCAAGCCAGCATAGCCAACTCCCGGTGAAACTCCGTCATAGATAACCTGCGGAG 1407
Db   :|||
QY 471 HisAsnAlaValSerSerTyrIle-----LysAspValPheArgValIleProSer 487
QY 1408 GAGGTGGTTGCAATTCATATGGCTCCAAAGGGAGCAGCGCAAGCGAGCTTAAG----- 1461
Db   :|||
QY 488 GluValLeuSerAsnSerTyrAsnLeuGlyGlnSerGlnValArgGlnLeuLysTyrGln 507
QY 1462 ---AACAAACACCCC 1473
Db   :|||
QY 508 GlyAsnSerGlyPro 512
```

RESULT 7

```
US-10-508-263-62
; Sequence 62, Application US/10508263
; Publication NO. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-508-263-62
```

```
Alignment Scores:
Pred. No.: 1,93e-58 Length: 472
Score: 779.50 Matches: 176
Percent Similarity: 51.6% Conservative: 88
Best Local Similarity: 34.4% Mismatches: 175
Query Match: 28.0% Indels: 10
DB: 9 Gaps: 10
```

US-10-728-323-3 (1-1524) x US-10-508-263-62 (1-472)

```
QY 4 CAGCAACCGGAGAG-----AACCGTGCCAGCTTCAGCGCCTCAATTCGCGAGACCT 57
Db   :|||
QY 27 GlnGlnGlyGlnGlnPheProAsnGluCysGlnLeuAspGlnLeuAsnAlaLeuGluPro 46
Db   :|||
QY 58 GACAATCCGATGATCAGAGCGGCTTACATCTGAGACTTCGAAACCCCAACCAACGAGAG 117
Db   :|||
QY 47 SerHisValLeuLysSerGluAlaGlyArgIleGluValTrpAspHisHisAlaProGln 66
QY 118 TTCGAATCGCGCGCGCTCTCTCGCTTAGTCTCCGCCGCCGCAACGCCCTTCGTAGG 177
Db   :|||
QY 67 LeuArgCysSerGlyValSerPheAlaArgTyrIleIleGluSerLysGlyLeuTyrLeu 86
QY 178 CCTTTCTACTCAATGCTCCCGAGAGATCTTCACAGCAAGGAGGAGGATACTTTGGG 237
Db   :|||
QY 87 ProSerPhePheAsnThrAlaLysLeuSerPheValAlaLysGlyArgGlyLeuMetGly 106
QY 238 TTGATATTCCTGCTGTTGCTTAGACACTATCAAGAGCCTCACACAAAGTCGTGCTGATCT 297
Db   :|||
QY 107 LysValIleProGlyCysAlaGluThrPheGlnAsp----- 118
QY 298 CAGTCCCAAGACCAACCAAGACGCTCTCAAGAGGAAGAACCAAGCAACAGCAACAGAT 357
Db   :|||
QY 119 SerSerGluPheGlnPro---ArgPheGluGlyGlnGlyGlnSerGlnArgPheArgAsp 137
QY 358 AGTCACCAAGAGGTGCACCGTTTCATGAGGTGATCTCATTGCGAGTTCCTCCACCGGTGT 417
Db   :|||
QY 138 MethHisGlnLysValGluHisIleArgSerGlyAspThrIleAlaThrThrProGlyVal 157
QY 418 GCTTTCGTGCTTACACACGACGACACTGATGTTGTTGCTTCTTCTTCTTACTCACACC 477
Db   :|||
QY 158 AlaGlnTrpPheTyrAsnAspGlyGlnGlnProLeuValIleValSerValPheAspLeu 177
QY 478 AACAAACCAACCAACCAAGTTCAGTTCCTCCAGGAGATTCATTTTGGCTGGGAACAG 537
Db   :|||
QY 178 AlaSerHisGlnAsnGlnLeuAspArgAsnProArgProPheTyrLeuAlaGlyAsn--- 196
QY 538 GAGCAAGAGTTCTTAAAGGTACCAACCAAGCAGACAAAGCAGACAAAGCAGAGCTTACCA 597
Db   :|||
QY 196 ----- 196
QY 598 TATAGCCCATACAGCCCGCAAGTCAGCTAGACAAAGAGCGGTGAATTTAGCCCTCGA 657
Db   :|||
QY 197 -----AsnProGlnGlyGlnValTrpLeuGlnGlyArgGlnGlnGlnProGln 212
QY 658 GGACAGCACACGCGCAGAGAACGAGCAGGACCAAGAAAGAAACCAAGGTGGAACATC 717
Db   :|||
QY 213 -----LysAsnIle 215
QY 718 TTCAGCGGCTTCAGCGGAGTTCTTGAAACAAGCTTTCAGGTTGACGACAGACAGATA 777
Db   :|||
QY 216 PheAsnGlyPheGlyProGluValIleAlaGlnAlaLeuLysIle---AspLeuGlnThr 234
QY 778 GTGCAAAACCTTAAGAGGCGACACCGAGAGTGAAGAGAGGGAGCCATTGTGACATGAGG 837
Db   :|||
QY 235 AlaGlnGlnLeu-----GlnAsnGlnAspAspAsnArgGlyAsnIleValArgValGln 252
QY 838 GGAGGCTTCAGAACTTCGAGCCCATAGAAAGAGACGTCGCGCAGCAAGAGAGGAATAC 897
Db   :|||
QY 253 GlyProPheGlyValIleArgProProLeuArgGlyGlnArgProGlnGlnGluGlu--- 271
QY 898 GATGAAGATGAATATGAATACATCGATGACAGAGATAGAAGCGGTGGCGGGAACGAGGC 957
Db   :|||
```

Db 272 -----GluGluGluGlyArgHisGlyArg----- 279
 QY 958 AGGGGAATGTTATTAAGAGAGCATCTGCACCGCAAGTGTCTAAAGAACATTTGGTAGA 1017
 Db 280 HisGlyAsnGlyLeuGluGluThrIleCysSerAlaArgCysThrAspAsnLeuAspAsp 299
 QY 1018 AACAGATCCCTCAGCATCTACAACTCTCAAGCTGCTACTCAAACTGCCAACGATCTC 1077
 Db 300 ProSerArgAlaAspValTyrIysProGlnLeuGlyTyrIleSerThrLeuAsnSerTyr 319
 QY 1078 AACCTTCTAATCTAGGTGGCTTGGACCTAGTGTCTGAATATGAAATCTCTACAGGAAT 1137
 Db 320 AspLeuProIleLeuArgPheIleArgLeuSerAlaLeuArgGlySerIleArgGlnAsn 339
 QY 1138 GCATTTTGTGCTCACTACACACCAACGACACAGCATCATATATCGATTGAGGGGA 1197
 Db 340 AlaMetValLeuProGlnTrpAsnAlaAsnAlaIleLeuTyrGluThrAspGly 359
 QY 1198 CGGGCTCAGCTGCAAGTCTGGAGACAGCAACGGCAACAGAGTGTAACGACGAGGAGCTTCAA 1257
 Db 360 GluAlaGlnIleGlnIleValAsnAspAsnGlyAsnArgValPheAspGlyGlnValSer 379
 QY 1258 GAGGGTCACGTGCTGTGGTCCACAGAACTTCGCCGTGCTGGAAAGTCCCAGAGCGAG 1317
 Db 380 GlnGlyGlnLeuIleAlaValProGlnGlyPheSerValValIysArgAlaThrSerAsn 399
 QY 1318 AACTTCGAATAGTGGCATTCAGACAGACTCAAGGCCCAGCATAGCCAACTCGCCGGT 1377
 Db 400 ArgPheGlnTrpValGluPheIysThrAsnAlaAsnAlaGlnIleAsnThrLeuAlaGly 419
 QY 1378 GAAAACTCCGTCTATAGTAACCTGCCGAGGAGGTGTTCGAAATTCATATGCCCTCCAA 1437
 Db 420 ArgThrSerValLeuArgGlyLeuProLeuGluValIleThrAsnGlyPheGlnIleSer 439
 QY 1438 AGGGAGCAGCAGGAGCTTAAGAACAAACACCCCTTCAAGTCTTCTCGTT----- 1488
 Db 440 ProGluAlaArgArgValIysPheAsnThrLeuGluThrThrLeuThrHisSerSer 459
 QY 1489 ---CCACGCTCTCAGCAGTCTCCGAGGGCTGTGGCT 1521
 Db 460 GlyProAlaSerTyrGlyArgProArgValAlaAla 471

RESULT 8

US-10-508-263-64
 ; Sequence 64, Application US/10508263
 ; Publication No. US20050260754A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BASF Plant Science GmbH
 ; TITLE OF INVENTION: Constructs and methods for regulating gene expression
 ; FILE REFERENCE: 53262-20085.00
 ; CURRENT APPLICATION NUMBER: US/10/508,263
 ; NUMBER OF SEQ ID NOS: 126
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 64
 ; LENGTH: 472
 ; TYPE: PR1
 ; ORGANISM: Arabidopsis thaliana
 US-10-508-263-64

Alignment Scores:

Pred. No.:	3 49e-58	Length:	472
Score:	776.50	Matches:	176
Percent Similarity:	51.6%	Conservative:	88
Best Local Similarity:	34.4%	Mismatches:	175
Query Match:	27.9%	Indels:	73
DB:	9	Gaps:	10

US-10-728-323-3 (1-1524) x US-10-508-263-64 (1-472)

QY 4 CAGCAACCGGAGGAG-----AACGGGTGCCAGTCTCCAGCGCTCAATGCGCAGACCT 57
 Db 27 GlnGlnGlyGlnGlnPheProAsnGluCysGlnLeuAspGlnLeuAsnAlaLeuGluPro 46

QY 58 GACAATCGCATTTGAATCAGAGGCGGTTACATTGAGACTTTGGAAACCCCAACACAGGAG 117
 Db 47 SerHisValLeuLeuSerGluAlaGlyArgIleGluValTrpAspHisHisAlaProGln 66
 QY 118 TTTCGAATGCGCGCGGTGCGCTCTCTCGCTTAGTCTCTCCGCCCAACCCCTTCGTAGG 177
 Db 67 LeuArgCysSerGlyValSerPheAlaArgTyrIleIleGluSerIysGlyLeuTyrLeu 86
 QY 178 CTTTCTTACTCCAACTGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGAGATCTTTGGG 237
 Db 87 ProSerPhePheAsnThrAlaIysLeuSerPheValAlaIaLysGlyArgGlyLeuMetGly 106
 QY 238 TTGATATTCCCTGGTGTCTAGACACTATGTAAGAGCCTCACACAAAGTCTGTCGATCT 297
 Db 107 LysValIleProGlyCysAlaGluThrPheGlnAsp----- 118
 QY 298 CAGTCCCAAGACCCACAGACGCTCTCCAAGGAGAGACCAAGCAACAGCAACAGAT 357
 Db 119 SerSerGluPheGlnPro---ArgPheGluGlyGlnGlyGlnSerGlnArgPheArgAsp 137
 QY 358 AGTCACCAAGAGGTGACCGTTTCGATGAGGTGATCTCATTCGAGTTCCTCCCGGTGT 417
 Db 138 MetHisGlnLysValGluHisIleArgSerGlyAspThrIleAlaThrThrProGlyVal 157
 QY 418 GCTTTTCTGCTCTACAACGACACACACTGATGTTGTTGCTGTTTCTCTTACTGACACC 477
 Db 158 AlaGlnTrpPheTyrAsnAspGlyGlnGluProLeuValIleValSerValPheAspLeu 177
 QY 478 AACAAACAACGACCAACAGCTTCAGTTCCCGCAGGAGATTCAAATTTGGCTGGGAACAG 537
 Db 178 AlaSerHisGlnAsnGlnLeuAspArgAsnProArgProPheTyrLeuAlaGlyAsn--- 196
 QY 538 GAGCAAGAGTCTTAAAGTACCAGCAACAAAGCAGACAAAGCAGACGAGAAAGCTTACCA 597
 Db 196 ----- 196
 QY 598 TATAGCCCATACAGCCCGCAAGTCAGCCTAGACAAGAGAGCGTGAATTTAGCCCTCGA 657
 Db 197 -----AsnProGlnGlyGlnValTrpLeuGlnGlyArgGlnGlnProGln 212
 QY 658 GGACAGCACAGCGCGCAGAGACGAGCAGGACGACAAAGAGAGAGAAACGAGGTGGAACATC 717
 Db 213 -----LysAsnIle 215
 QY 718 TTCAGCGGCTTCACGCCGAGTTCCTGGAAACAGCCTTCAGGTTCCAGGTTGACGACAGATA 777
 Db 216 PheAsnGlyPheGlyProGluValIleAlaGlnAlaLeuIysIle---AspLeuGlnThr 234
 QY 778 GTGCAAAACCTAAGAGCGCAGACCGAGAGTGAAGAGAGGAGCGCATTTGTGACAGTGAGG 837
 Db 235 AlaGlnGlnLeu-----GlnAsnGlnAspAsnArgGlyAsnIleValArgValGln 252
 QY 838 GGAGGCTTCAGAAATCTTGAGCCGACATAGAAAAGAGACGTGCCGACGAGAAGAGGAAATAC 897
 Db 253 GlyProPheGlyValIleArgProProLeuArgGlyGlnArgProGlnGluGlu--- 271
 QY 898 GATGAAGATGAATATGAATACGATGAAGAGATAGAGCGGTGGCGGGGAGGACGAGGC 957
 Db 272 -----GluGluGluGlyArgHisGlyArg----- 279
 QY 958 AGGGGAATGTTATTAAGACAGACGATCTGCACCGCAAGTGTCTAAAGAACATTTGGTAGA 1017
 Db 280 HisGlyAsnGlyLeuGluGluThrIleCysSerAlaArgCysThrAspAsnLeuAspAsp 299
 QY 1018 AACAGATCCCTCAGCATCTACAACTCTCAAGCTGCTACTCAAACTGCCAACGATCTC 1077
 Db 300 ProSerArgAlaAspValTyrIysProGlnLeuGlyTyrIleSerThrLeuAsnSerTyr 319
 QY 1078 AACCTTCTAATCTAGGTGGCTTGGACCTAGTGTCTGAATATGAAATCTCTACAGGAAT 1137
 Db 320 AspLeuProIleLeuArgPheIleArgLeuSerAlaLeuArgGlySerIleArgGlnAsn 339

Db 412 AsnPheAlaValIleLysArgAlaAsnGluGlnGlySerArgTrpValSerPheLysThr 431
QY 1345 GACTCAAGGCCAGCATAGCAACCTCGCGGTGAATACTCGTCTATAGATCACTGCGG 1404
Db 432 AsnAspAlaMetIleAlaAsnLeuAlaGlyArgValSerAlaSerAlaSerPro 451
QY 1405 GAGGAGGTGTCCTCAATTCATATGCTCCCAAGGAGCAGGACGAGGCTTAAAG 1461
Db 452 LeuThrLeuTrpAlaAsnArgfyrGlnLeuSerArgGluGluAlaGlnGlnLeuLys 470

RESULT 10

US-10-508-263-10
; Sequence 10, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Brassica napus
US-10-508-263-10

Alignment Scores:
Pred. No.: 5,15e-55 Length: 490
Score: 739.50 Matches: 166
Percent Similarity: 49.5% Conservative: 86
Best Local Similarity: 32.6% Mismatches: 162
Query Match: 26.6% Indels: 95
DB: 9 Gaps: 7

US-10-728-323-3 (1-1524) x US-10-508-263-10 (1-490)

QY 19 AACGGTGCAGTTCCAGCGCCTCAATGGCGACAGACCTGACAACTGCATTCGATTCAGAG 78
Db 28 AsnGluCysGlnLeuAspGlnLeuAsnAlaLeuGluProSerHisValLeuLysAlaGlu 47
QY 79 GCGGGTTACATTCAGACTTGAACCCCAACACACAGGAGTTCGAATGCGCGCGTCGCC 138
Db 48 AlaGlyArgIleValTrpAspHisHisAlaProGlnLeuArgCysSerGlyValSer 67
QY 139 CTCTCTCGCTTAGCTCTCCGCGCAACGCCCTTCGTAGCGCTTCTACTCCAATGCTCCC 198
Db 68 PheValArgfyrIleIleGluSerLysGlyLeuTyLeuProSerPhePheSerThrAla 87
QY 199 CAGGAGATCTTCATCCAGCAGGAAGGGATACCTTTGGTTGATATATCCCTGTTGCTCT 258
Db 88 LysLeuSerPheValAlaLysGlyGlyLeuMetGlyArgValProGlyCysAla 107
QY 259 AGACACTATGAACAGCTCACACA----- 282
Db 108 GluThrPheGlnAspSerSerValPheGlnProSerGlyGlySerProSerGlyGlyGly 127
QY 283 CAAGGTGCTGCATCTCAGTCCCAAGACACCAAGACAGCTCTCAAGGAGAAGACCAAGC 342
Db 128 GlnGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 147
QY 343 CAACAGCAA-----CGAGATAGTACCAG 366
Db 148 GlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 167
QY 367 AAGGTGACCGTTTCGATGAGGGTGATCTCATTCAGTCCCGAGTCCCGCGGTGTTCTTGG 426
Db 168 LysValGluHisIleArgThrGlyAspThrIleAlaThrHisProGlyValAlaGlnTrp 187
QY 427 CTCTACAACGACCAACGACATGATGTTGCTGTTCTTCTTACTGACCAACAACAAC 486
Db 188 PheTyAsnAspGlyAsnGlnProLeuValIleValSerValLeuAspLeuAlaSerHis 207

RESULT 11

US-10-508-263-18
; Sequence 18, Application US/10508263
; Publication No. US20050260754A1

QY 487 GACACCAAGCTTGTAGTCCCGAGGAGATTCAATTTGGCTGGCAACACGAGCAAGAG 546
Db 208 GlnAsnGlnLeuAspArgAsnProArgProPheTyLeuAlaGlyAsn----- 223
QY 547 TTCTTAAGGTACCAACAAGCAGACAAAGCAGACAAAGCAGCAAGAGCTTACCATATAGCCCA 606
Db 223 ----- 223
QY 607 TACAGCCGCGCAAGTTCAGCTTAGACAAGAAGACGTCGATTTAGCCCTCGAGGACAGCAC 666
Db 224 ---AsnProGlnGlyGlnValTrpIleGluGlyArgGluGlnGlnProGln----- 239
QY 667 AGCCCGCAGACGACGAGCAGGACAAGAAGAAAGAGGTTGGGAAACATCTTCAGCCGC 726
Db 240 -----LysAsnIleLeuAsnGly 245
QY 727 TTCAGCCGCGAGTTCCTGGAACAACGCTTCAGAGTTGACGACAGAG-----CAGATAGTG 780
Db 246 PheThrProGluValLeuAlaLysAlaPheLysIleAspValArgThrAlaGlnGlnLeu 265
QY 781 CAAAACCTTAAGAGCGACACCGAGAGTGAAGAAGAGGAGGCCATTTGTGACAGTGAAGGA 840
Db 266 GlnAsn-----GlnGlnAspAsnArgGlyAsnIleIleArgValGlnGly 280
QY 841 GGCCTCAGNACTTCAGGCCCATAGAAAGAGACGTCGCGCAGCAAGAGAGGAATACGAT 900
Db 281 ProPheSerValIleArgProLeuArgSerGlnArgProGlnGlnThrGluVal---- 299
QY 901 GAAAGATGAATATGAAATACGATGAAGAGTAGAAGCGCTGGCAGGGAAGCAGAGCAGG 960
Db 299 ----- 299
QY 961 GGGAAATGTTATGAAGAGACGATCTGCACCGCAAGTGTAAAGAAACAATTCGTTAGAAAC 1020
Db 300 ---AsnGlyLeuGluGluThrIleCysSerAlaArgCysThrAspAsnLeuAspAspPro 318
QY 1021 AGATCCCTGACATCTACAACCTCAAGCTGCTCACTCAAAACCTGCCAACCATCTCTCAAC 1080
Db 319 SerAsnAlaAspValTyLysProGlnLeuGlyTyIleSerThrLeuAsnSerTyAsp 338
QY 1081 CTCTTAATACTTGTAGTGGCTTGGACCTAGTCTGTAATATGAAATCTCTACAGGAATGCA 1140
Db 339 LeuProIleLeuArgPheLeuArgLeuSerAlaLeuArgGlySerIleArgGlnAsnAla 358
QY 1141 TTGTTTGTGCTACTACAAACACGACACGACATCATATCGATTGAGGGGACCG 1200
Db 359 MetValLeuProGlnTrpAsnAlaAsnAlaValLeuTyxValThrAspGlyGlu 378
QY 1201 GCTCAGTCGACAGTCTGGACAGCAACGCGACAGAGTGTACGAGGAGGCTTCAAGAG 1260
Db 379 AlaHisValGlnValValAsnAspAsnGlyAspArgValPheAspGlyGlnValSerGln 398
QY 1261 GGTCACTGCTTGTGTGTCACAGAACTTCGCGCTGCTGGAAAGTCCACAGGCGAGAAC 1320
Db 399 GlyGlnLeuLeuSerIleProGlnGlyPheSerValValLysArgAlaThrSerGluGln 418
QY 1321 TTCGAATACGTCGATTCGAAGACAGACTCAAGGCCCGCATAGCCAACTCCCGCGTGAA 1380
Db 419 PheArgTrpIleGluPheLysThrAsnAlaAsnAlaGlnIleAsnThrLeuAlaGlyArg 438
QY 1381 AACTCGTCATAGATAACCTCGCGAGGAGGTGTTGCAAAATTCATATCGCTCCCAAGG 1440
Db 439 ThrSerValLeuArgGlyLeuProLeuGluValIleSerAsnGlyTyIleGlnIleSerLeu 458
QY 1441 GAGCAGGCAAGGACGCTTAAGAACACAC 1467
Db 459 GluGluAlaArgArgValLysPheAsn 467

```

; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Brassica napus
US-10-508-263-18

Alignment Scores:
Pred. No.: 5,21e-55 Length: 509
Score: 739.50 Matches: 164
Percent Similarity: 49.9% Conservative: 91
Best Local Similarity: 32.1% Mismatches: 169
Query Match: 26.6% Indels: 87
DB: 9 Gaps: 8

US-10-728-323-3 (1-1524) x US-10-508-263-18 (1-509)
QY 19 AACGGTGCAGTTCCAGCGCTCAATGCGCAGACCTGACACATCGCATGAATCAGAG 78
DB 35 AsnAlaCysAsnLeuAspValLeuGlnProThrGluThrIleLysSerGlu 54
QY 79 GCGGTTACATTCAGACTTGGAAACCCCAACACACAGGAGTTTCGATGCGCGCTGCC 138
DB 55 AlaGlyArgValGluTrpAspHisAsnAsnProGlnIleArgCysAlaGlyValSer 74
QY 139 CTCTCTCGTTCAGTCTCCGCGCAACGCGCTTCGTAGCGCTTCTTACTCCCAATGCTCC 198
DB 75 ValSerArgValIleIleGlnGlyGlyLeuThrLeuProThrPhePheSerSerPro 94
QY 199 CAGGAGATCTTCATCCAGCAAGGAAGGGATCTTTGGGTTCATATCCCTGGTGTCTCT 258
DB 95 LysIleSer-Tyr-ValValGlnGlyMetGlyIleSerGlyArgValValProGlyCysAla 114
QY 259 AGACACTAT-----GAAGAGCT-----CACACACAGGTGCTGCATCTCAGTCC 303
DB 115 GluThrPheMetAspSerGlnProMetGlnGlyGlnGlnGlyGlnProTrpGlnGly 134
QY 304 CAAGACCAACCAAGAGCTCTCAAGGAGAA-----GACCAAGCCAAACAGCAA-----CGA 354
DB 135 GlnGlnGlyGlnGlnGlyGlnGlnGlyGlnGlnGlyGlnGlnGlyGlnGlnGlnGln 154
QY 334 -----GACCAAGCCAAACAGCAA-----CGA 354
DB 155 GlyGlnGlnGlyGlnGlnGlyGlnGlnGlyGlnGlnGlyGlnGlnGlyPheArg 174
QY 355 GATAGTCACCAAGGTGCACCGTTTCGATGAGGTGATCTCATTTGCAGTCCACCGGT 414
DB 175 AspMetHisGlnLysValGluHisValArgHisGlyAspIleAlaIleThrAlaGly 194
QY 415 GTTGCTTTCTGCTCTACCAACCAACCAACCACTGATGTTGTTGTTCTTCTTACTGAC 474
DB 195 SerSerHisTrpIleTyrAsnThrGlyAspGlnProLeuValIleIleCysLeuLeuAsp 214
QY 475 ACCAACAACAACCAACCAACCACTGATGATTTCCCGAGGAGATTCATATTTGGCTGGGAAC 534
DB 215 IleAlaAsnTyr-GlnAsnGlnLeuAspArgAsnProArgThrPheArgLeuAlaGlyAsn 234
QY 535 ACGGAGCAGAGTTCTTAGGTACCGCAACCAACCAACCAAGCAGACCAAGAGAGCTTCA 594
DB 235 AsnProGlnGlyGlySerGlnGlnGlnGlnGlnGln----- 248
QY 595 CCATATAGCCCATACAGCCCGCAAGTCAGCTAGACAGAAGAGCGTGAATTTAGCCCT 654
DB 248 ----- 248
QY 655 CGAGGACAGCACAGCGCCGACAGACGAGCAGGACAGAGAGAGAGAGAGAGAGAGAGAG 714

```

```

Db 249 -----Asn 249
QY 715 ATCTTCAGCGCTTCACCGCGAGTTCCTGGAAACAGCCTTCCAGGTTGACGACAGACAG 774
Db 250 MetLeuSerGlyPheAspProGlnValLeuAlaLeuLeuValIleAspValArg--- 268
QY 775 ATAGTCAAAAACCTAAGAGCGCAGACCGAGAGTGAAGAAGAGGAGCCATTGTGACAGTG 834
Db 269 LeuAlaGlnGluLeuGlnAsnGlnAspSer-----ArgGlyAsnIleValArgVal 286
QY 835 AGGGAGGCGCTCAGAAATCTTGAGCCACGATGAGAAGAGACGTGCGCAGCAAGAGAGAA 894
Db 287 LysGlyProPheGlnValValArgProLeuArgGln----- 299
QY 895 TACGATGAAGATGAATATGATATGATGATGATGATGATGATGATGATGATGATGATGATG 954
Db 300 -----ProTyrGluSerGluGlnTrpArgHisProArgGlyProProGlnSer 315
QY 955 GCGAGGGGAATGTTATTAAGAGACGATCTGACCGCAAGTGTCTAAAGAAACATTTGTT 1014
Db 316 ProGlnAspAsnGlyLeuGluGluThrIleCysSerMetArgThrHisGluAsnIleAsp 335
QY 1015 AGAAACAGATCCCTGACATCTAACCCCTCAACCTCAACCTGTTCTCACTCAAACTCCCAACGAT 1074
Db 336 AspProAlaArgAlaAspValTyrLysProAsnLeuGlyArgValThrSerValAsnSer 355
QY 1075 CTCACACTTCTAATACTTGTAGTGGCTTGACCTAGTGTGTAATATGGAATCTCTACAGG 1134
Db 356 TyrThrLeuProIleLeuGlnIleArgLeuSerAlaThrArgGlyIleLeuGlnGly 375
QY 1135 AATGATTGTTTGTGCTCACTACAAACCAACGACACACATCATATATCGATTGAGG 1194
Db 376 AsnAlaMetValLeuProLysTyrAsnMetAsnAlaAsnGluIleLeuTyrCysThrGln 395
QY 1195 GGACGGCTCAGTGCAGTCAAGTGTGACAGCAACGCGCAACAGAGTGTACGACGAGAGCTT 1254
Db 396 GlyGlnAlaArgIleGlnValValAsnAspAsnGlyGlnAsnValLeuAspGlnGlnVal 415
QY 1255 CAAGAGGTGCTGCTGTTGTGTCGACAGAACTTTCGCGCTGCTGCGTGAAGTCCAGAGC 1314
Db 416 GlnLysGlyGlnLeuValIleProGlnGlyPheAlaTyrValValGlnSerHisGln 435
QY 1315 GAGAACTTCGATACGTGCGCATTCAGACAGACTCAAGGCCCGACGATAGCCAACTCGCC 1374
Db 436 AsnAsnPheGluTrpIleSerPheLysThrAsnAlaAsnAlaMetValSerThrLeuAla 455
QY 1375 GGTGAAAACCTCGCTCATAGATAAATCTGCGGAGGAGGTGTTGCAAAATTCATATGCGCTC 1434
Db 456 GlyArgThrSerAlaLeuArgAlaLeuProLeuGluValIleThrAsnAlaPheGlnIle 475
QY 1435 CAAGGGAGCAGGCAAGGCGAGCTTAAGAACAC 1467
Db 476 SerLeuGluGluAlaArgArgIleLysPheAsn 486

RESULT 12
US-10-508-263-66
; Sequence 66, Application US/10508263
; Publication No. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-508-263-66

```

Alignment Scores:
 Pred. No.: 1,34e-54 Length: 455
 Score: 734.50 Matches: 171
 Percent Similarity: 47.9% Conservative: 73
 Best Local Similarity: 33.6% Mismatches: 174
 Query Match: 26.4% Indels: 91
 DB: 9 Gaps: 8

US-10-728-323-3 (1-1524) x US-10-508-263-66 (1-455)

QY	19	AACGGCTCCAGTTCACGCGCTCAATGCGCAGACCTTGACAATCGCATTCGAATTCAGATCAGAG	78
		:::	
DB	28	AsnGluCysGlnLeuAspGlnLeuAsnAlaLeuGluProSerGlnIleIleIysSerGlu	47
QY	79	GGCGGTTACATTGAGACTTGGAAACCCCAACACAGGAGTTCGAATGCGCGCGGTGCGC	138
		:::	
DB	48	GlyGlyArgIleGluValTrpAspHisHisAlaProGlnLeuArgCysSerGlyPheAla	67
QY	139	CTCTCTGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCGCTTCTACTCCAAATGCTCC	198
		:::	
DB	68	PheGluArgPheValIleGluProGlnGlyLeuPheLeuProThrPheLeuAsnAlaGly	87
QY	199	CAGGAGATCTTCATCCACCAAGGAAGGGATACITTTGGGTGTGATATTCCTCGTTGTCT	258
		:::	
DB	88	LysLeuThrPheValValHisGlyArgGlyLeuMetGlyArgValIleProGlyCysAla	107
QY	259	AGACACTATGAAGACCTCACACAAAGTCTGTCATCTCAGTCCCAAGACCAACAAGA	318
DB	108	GluThrPheMetGluSerProValPheGlyGluGly	119
QY	319	CGTCTCCAAGAGAGAACCAAGCCAACAGCAGATAGTCACCAAGGTGCACCGT	378
		:::	
DB	120	-----GlnGlyGlnSerGlnGlyPheArgAspMetHisGlnIysValGluHis	137
QY	379	TTCGATGAGGTGATCTCATTCGAGTTCACCGGTGTGCTTTCTGGCTCTACAAACGAC	438
		:::	
DB	138	LeuArgCysGlyAspThrIleAlaThrProSerGlyValAlaGlnTrpPheTrpAsnAsn	157
QY	439	CACGACACTGATGTTGGTGTCTTCTTACTGACACCAACAACACCAACACAGCTT	498
		:::	
DB	158	GlyAsnGluProLeuIleLeuValAlaAlaAspLeuAlaSerAsnGlnAsnGlnLeu	177
QY	499	GATCAGTTCCTCCAGGAGATTCATTTGGCTGGGAACACGGAG-----CAAGAGTTCTT	552
		:::	
DB	178	AspArgAsnLeuArgProPheLeuIleAlaGlyAsnAsnProGlnGlyGlnGluTrpLeu	197
QY	553	AGGTACACGACAAAGACAGACAAAGACGACCAAGAGCTTACCATATAGCCCATACACG	612
		:::	
DB	198	GlnGlyArgIysGlnGlnIysGlnAsn-----	206
QY	613	CCGCAAGTTCACCTAGACAAGAAGCGTGAAATTTAGCCCTCGAGGACGACACGCGC	672
DB	206	-----	206
QY	673	AGAGAACGACGAGCAACAGAAACGAAAGTGGAAACATCTTCAGCGCTTCACG	732
DB	207	-----AsnIlePheAsnGlyPheAla	213
QY	733	CCGGAGTTCCTGGAAACAAGCCCTTCAG-----GTTGACGACAGACAGATAGTGC	786
		:::	
DB	214	ProGluIleLeuAlaGlnAlaPheIysIleAsnValGluThrAlaGlnGlnLeuGlnAsn	233
QY	787	CTAAGAGCGGACCGAGAGTGAAGAAGGAGGCCATTTGTACAGTGTAGGGGAGCGCTC	846
		:::	
DB	234	-----GlnGlnAspAsnArgGlyAsnIleValIysValAsnGlyProPhe	248
QY	847	AGAACTTTGAGCCAGATAGAAGAAGACGTGCCGCAAGAAAGAAATCATGATGAAGT	906
		:::	
DB	249	GlyValIleArgPro-----	253
QY	907	GAATATGAATACGATGAAGAGGATAGAAGCGCTGCGAGGGAAGCAGAGCCAGG-----	960
		:::	
DB	254	-----ProLeuArgArgGlyGluGlnGlnProHisGlu	266


```
Db      108  GluThrPheGlnAspSerValPheGlnProGlySerGlySerProPheGlyGluGly 127
QY      277  CACACACAAGTCGTCGATCTCAG-----TCC 303
Db      128  GlnGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 147
QY      304  CAAAGACCACCAAGACGCTCCAAAGGAGAAGACCAAGGACCAACAGCAA----- 351
Db      148  GlnGlyGlnGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 167
QY      352  ---CGAGATAGTCACACAGAGGTGCACCCGTTTCGATGAGGGGTGATCTCAATTGCGTCCC 408
Db      168  PheArgAspMetHisGlnLysValGluHisIleArgSerGlyAspThrIleAlaThrHis 187
QY      409  ACCGGTGTTCCTTTCGCTCTACAAAGACGACGACATGATGTGTGCTGCTTCTCTTT 468
Db      188  ProGlyValAlaGlnTrpPheTyrAsnAsnGlyAsnGlnProLeuValIleValAlaVal 207
QY      469  ACTGACACCAACAACACACACAGCTTGATCAGTTCCTCCAGGAGATTCAATTTGGCT 528
Db      208  MetAspLeuAlaSerHisGlnAsnGlnLeuAspArgAsnProSerGlnPheTyrLeuAla 227
QY      529  GGAACACACGAGCAAGATTCTTAAGGTACCAGCAACAAGACAGACAAAGACGACGAAGA 588
Db      228  GlyLys----- 229
QY      589  AGCTTACCATATAGCCCATACAGCCCGCAAGTCAGCTAGACAAAGACGAGTGTGATTT 648
Db      229  ----- 229
QY      649  AGCCCTCGAGGACAG-----CACAGCGCAGAGACGAGCAGCAAGAAGAAGAA 699
Db      230  AsnProGlnGlnSerTrpLeuHisGlyArg-----GlyGlnGlnProGln 245
QY      700  AACGAAGGTGGAAACATCTTCAGCGCTTCACGCCGAGTTCCTCGAAACAAGCTTCCAG 759
Db      246  Asn-----AsnIleLeuAsnGlyPheSerProGluValLeuAlaGlnAlaPheLys 262
QY      760  GTTGACGACAGA-----CAGATAGTGCAAAACCTAAGAGCGCAGACCGAGAGTGAAGAA 813
Db      263  IleAspValArgThrAlaGlnGlnLeuGlnAsn-----GlnGlnAspAsn 277
QY      814  GAGGAGGACATCTGACAGTGGGAGGCGCTCAGAACTCTTGAGCCCGACATAGAAAGAGA 873
Db      278  ArgGlyAsnIleValArgValGlnGlyProPheGlyValIleArgProLeuLysSer 297
QY      874  CGTCCGACGAAGAAGGAATACGATGAAGATGAATATGAATACGATGAAGAGGATAGA 933
Db      298  GlnArgProGlnGluThrGlu----- 304
QY      934  AGCGGTGGCAGGGAAGCAGAGCAGGGGGAATGTTATTGAAGACAGATCTGCACCGCA 993
Db      305  -----AlaAsnGlyLeuGluGluThrIleCysSerAla 315
QY      994  AGTGCTAAAAAAGAACATTGGTAGAAACAGATCCCTGACATCTACAAACCTCAAGCTGGT 1053
Db      316  ArgCysThrAspAsnLeuAspAspProSerAsnAlaAspValTyrLysProGlnLeuGly 335
QY      1054  TCACCTCAAACTGCCACGATCTCAACCTTCTAATACTTAGTGGCTGGACCTAGTGT 1113
Db      336  TyrIleSerIleLeuAsnSerTyrAspLeuProIleLeuArgValLeuArgLeuSerAla 355
QY      1114  GAATATGGAATCTCTACAGGAATGATTTGTTGCTCCTCACTACACACCAACGACACAC 1173
Db      356  LeuArgGlySerIleArgGlnAsnAlaMetValLeuProGlnTrpLysSerTyrAsn 375
QY      1174  AGCATCATATCGATTGAGGGACGGCTCAGTGCAGTCTCGTGACGACGACGACCAAC 1233
Db      376  AlaValPheAspGlyValThrAspGlyGluAlaGlnIleGlnValValAsnAspGlnLys 395
QY      1234  AGAGTGTACGACGAGAGCTTCAAGAGGGTTCAGCTGTGTGTGTCGACAGAACTTCGCC 1293
Db      396  ArgValPheAspGlyGlnValSerGlnGlnLeuLeuSerIleProGlnGlyPheSer 415
```

```
QY      1294  GTCGCTGGAAGTCCCGAGCGAGAACTTCGATACGTGGCATTCAAGACAGACTCAAGG 1353
Db      416  ValValLysArgAlaThrSerAspGlnPheArgTrpIleGluPheLysThrAsnAlaAsn 435
QY      1354  CCCAGCATAGCCCAACTCCCGCGTGAAAACTCCGTCATAGATAAACCCTGCCGAGGAGGTG 1413
Db      436  AlaGlnIleAsnThrLeuAlaGlyArgThrSerValMetArgGlyLeuProLeuGluVal 455
QY      1414  GTTCAAATTCATATGAGCTCCAAAGGAGGACGAGGACGACCTTAAGAACAAAC 1467
Db      456  IleAlaAsnGlyTyrGlnIleSerLeuGluGluAlaArgArgValLysPheAsn 473
```

RESULT 15

```
US-10-508-263-68
; Sequence 68, Application US/10508263
; Publication NO. US20050260754A1
; GENERAL INFORMATION:
; APPLICANT: BASF Plant Science GmbH
; TITLE OF INVENTION: Constructs and methods for regulating gene expression
; FILE REFERENCE: 53262-20085.00
; CURRENT APPLICATION NUMBER: US/10/508,263
; CURRENT FILING DATE: 2004-09-20
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 68
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-508-263-68
```

```
Alignment Scores:
Pred. No.:          9,4e-51      Length:      451
Score:              689.50      Matches:    157
Percent Similarity: 48.0%      Conservative: 85
Best Local Similarity: 31.2%   Mismatches: 174
Query Match:        24.8%      Indels:     88
DB:                  8
Gaps:                9

US-10-728-323-3 (1-1524) x US-10-508-263-68 (1-451)
```

```
QY      1  CGCGACGCAACCCGAGGAGAACGCGTCCAGTTTCCAGCGCTCAATGCCGAGACCTGAC 60
Db      28  ArgGluAlaProPheProAsnAlaCysHisPheSerGlnIleAsnSerLeuAlaProAla 47
QY      61  AATCGCATTAATCATGAGCGCGTTACATTGAGACTTGGAACCCCAACACGAGGAGTTC 120
Db      48  GlnAlaThrLysPheGluAlaGlyGlnMetGluValTrpAspHisMetSerProGluLeu 67
QY      121  GAATGCGCGCGTCCGCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCGCT 180
Db      68  ArgCysAlaGlyValThrValAlaArgIleThrLeuGlnProAsnSerIlePheLeuPro 87
QY      181  TTCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGGGATACTTTGGGTG 240
Db      88  AlaPhePheSerProProAlaLeuAlaTyrValValGlnGlyGlyValMetGlyThr 107
QY      241  ATATTCCTCGTGTTCCTAGACATATGAAGACCTCACACACAGGTGCTCGATCTCAG 300
Db      108  IleAlaSerGlyCysProGluThrPheAlaGluValGluGlySerSerGlyArgGlyGly 127
QY      301  TCCCAAGACCAACCAACGAGCTCTCCAAAGGAGAAGACCAAAAGCAGCAACGAGATAGT 360
Db      128  GlyGlyAspProGlyArgArgPheGlu-----AspMet 138
QY      361  CACCAAGGTGCACCGTTTCGATGAGGGTGATCTCATTCGATTCCTCCCGGTGTGCT 420
Db      139  HisGlnLysLeuGluAsnPheArgArgGlyAspValPheAlaSerLeuAlaGlyValSer 158
QY      421  TTCTGGCTCTACACGACCAACGACACTGATGTGTGCTGTTCTCTTACTGACACCAAC 480
Db      159  GlnTrpTrpTyrAsnArgGlyAspSerAspAlaValIleValLeuAspValThr 178
```


THIS PAGE LEFT BLANK

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame plus n2p model

Run on: May 15, 2006, 21:46:12 ; Search time 10.2004 Seconds
(without alignments)
4312.597 Million cell upd

Title: US-10-728-323-3

Perfect score:

Sequence: 1 cggcagcaaccggaggagaa.....ctccgagggctgtggcttaa 1524

Scoring table:

Xgapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Fgapop	6.0	Fgapext	7.0
Delop	6.0	Delext	7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Ma

Post-processing: Minimum Match 0% Maximum Match 100%

Maximum match 100%
Listing first 45 summaries

Command line parameters:

```

COMMUNITYLINE PARAMETERS.
-MODEL=frame+ n2p.model -DEV=xl1p
-Q=/abs/ABSSWE.spool/USU1728323/rumat_15052006_172134_22413/app_query.fasta_1
-DB=PIR -QPMT=fastaan -SUFFIX=n2p.tpr -MINMATCH=0.1 -LOOPEI=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bloem62 -TRANS=human40.15 -LIST=45
-DOCALIGN=700 -THAT SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=1 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs02p
-USER=USU1728323 @CGN_1_92 @rumat_15052006_172134_22413 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

```
Database : PIR_80:*
1: pir1
2: pir2
3: pir3
4: pir4
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	1502	54.0	481	2	S04605	glycinin G3 - soyb
2	1499.5	54.0	485	2	S11002	glycinin G2 precu
3	1496	53.8	495	2	S10851	glycinin G1 precu
4	1492.5	53.7	485	1	FWSVG1	glycinin chain A2B
5	1466	52.8	495	1	FWSVG2	glycinin chain A1a
6	1457.5	52.4	498	2	S44294	legumin A2 precuso
7	1455	52.4	520	2	S08237	legumin A2 precuso
8	1440	51.8	517	1	FWPMLA	legumin A precuso
9	1431.5	51.5	482	2	S49877	legumin A precuso
10	1428	51.4	484	2	S11003	glycinin G3 precu
11	1414.5	50.9	500	2	S14393	legumin A2 precurs
12	1384	49.8	507	2	T06452	probable legumin A
13	1375	49.5	497	2	S14392	legumin A1 precurs
14	1073.5	38.6	551	2	S51941	prunin 1 precursor

ALIGNMENTS

RESULT 1

REF ID: A5301
S04605

glycinin G3 - soybean

C:Species: Glycine max (soybean)
C:Date: 28-Feb-1990 #sequence revision 28-Feb-1990 #text change 05-Oct-2004

C:Accession: S04605

R:Cho, T.J.: Nielsen, N.C.

Nucleic Acids Res. 17, 4388, 1989

A;Title: The glycinin Gy(3) gene from soybean.

A;Reference number: S04

A;Accession: S04605

A;Molecule type: DNA

A;Residues: 1-481 <CHO>

A; Cross-references: UNI:

A; Experimental source:

C;Genetics:

Alignment Scores:

Pedigree No.	1.34e-103	Length:	481
Score:	1502.00	Matches:	297
Percent Similarity:	71.6%	Conservative:	68
Best Local Similarity:	58.2%	Mismatches:	91
Query Match:	54.0%	Indels:	54
DB:	2	Gaps:	7

US-10-728-323-3 (1-1524) x 504605 (1-481)

Qy	1	CGGCACCAACCGGAGGAGAACCGGCGCCAGTCTCAGCGCCTCAATGCGCAGAGACCTTAC	60
Db	23	ArggluInPrGlnGlnAsnGluCysGlnIleGlnArgLeuAsnIleuLeuArgProAsp	42
Qy	61	AATCGATTGAATCAGAGGGCGGTTACATTAGACTTCGGAACCCCAACAACGAGGAGTTC	120
Db	43	AsnArgIleGluSerGluGlyPheIleGluThrIrpAsnProAsnAsnIleProPhe	62
Qy	121	GAATGCGCCGGGGTGGCCCTCTCTCGCTTAGTCTCTCCGCGCAACGCCCTCGTAGGCCT	180

```

Db      63  GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 82
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      181  TTCTACTCCAAATCCTCCACAGAGATCTTCATCCACAGAGGAGGCGATACTTTGGTTG 240
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      83  SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 102
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      241  ATATTCCCTGGTTGCTTAGACACTATGAAGACCTTCACACAGAGTGTGTCATCTCAG 300
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      103  IlePheProGlyCysProSerThrPheGluGluProGlnGlnGlySGLY-----Gln 119
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      301  TCCCAAGACCAACCAAGCGTCTCCAGGAGAGACCAAGCCAAACAGCAAGAGATAGT 360
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      120  SerSerArgPro-----GlnAspArg 126
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      361  CACCAAGAGTGCACCGTTTCGATGAGGGTGATCTCATTCAGTTCCTCCACCGGTGTGCT 420
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      127  HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla 146
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      421  TTCTGGCTCTAACACACGACGACTGATGTTGTTGCTGTTCTTCTTACTGACCAAC 480
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      147  TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn 166
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      481  AACCAACGACACACAGCTTGATCAGTTCCCGCAGGAGATTCAATTTGGCTGGACACGGAG 540
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      167  SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      541  CAAGAGTTCTTAAAGTACCAGCAACAAAGACAGACAGACAGACAGAAAGACTTACCATAT 600
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      187  GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly-----199
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      601  AGCCCATACAGCCCGCAAGTCAAGCTAGACAAAGAGAGCGGTGAATTGACCTCGAGGA 660
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      200  -----Gly 200
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      661  CACACAGCGCAGAGAACGACGACGACAGAGAGAGAGAAACGAAGTGGAAACATCTTC 720
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      201  ThrGlnSerGlnLysGlyLysArgGlnGlnGlnGluGlnGluAsnGluGlySerIleLeu 220
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      721  AGCGGTTTCACGCGCGGAGTTCTCTGGAAACAAAGCCCTCCAGGTTGACGACAGACAGATAGT 780
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      221  SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal 239
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      781  CAAACCTAAGAGCGCAGACCGAGAGTGAAGAGAGGAGCCATTGTGACAGTGGAGGA 840
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      240  ArgLysLeuGlnGlyGluAsnGluGluGluGluLysGlyAlaIleValThrValLysGly 259
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      841  GGCTCAGAAATCTTGAGCCCA-----GATAGAAAGAGACGTGCGCAGCAAGAGAG 891
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      260  GlyLeuSerValIleSerProProThrGluGluGlnGlnGlnArgProGluGluGluGlu 279
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      892  GAATACGATGAAGATGAATATGAATACGATGAAGAGATAGAAAGCGTGGCAGGGGAAGC 951
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      280  LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer 294
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      952  AGAGGAGGGGGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTCTAAAGAACACATT 1011
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      295  Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle 311
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      1012  GGTTAGAAACAGATCCCTGCATCTCAACACCTCAAGCTGGTTGTTCACTCAAAACCTGCCAAC 1071
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      312  GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr 331
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      1072  GATCTCAACACTTCTAATCTAGGTGGCTTGGACCTAGTGTGCTGAATATGAAATCTCTAC 1131
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      332  SerLeuAspPheProAlaLeuSerTyrPheLeuLysLeuSerAlaGlnPheGlySerLeuArg 351
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      1132  AGGAATGCATTGTTGTGCTCTACTACCAACCAACGACACAGCATCATATATCATTTG 1191
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      352  LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu 371
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      1192  AGGGGACGGGCTCACGTGCAAGTCTGTGGACAGCAACGGCAACAGAGTGTACGACGAGGAG 1251
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

```

```

Db      372  AsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGlyGlu 391
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      1252  CTTCAAGAGGGTCACGTGCTTGTGTGTCACAGAACTTCGCGCTGCGTGAAGTCCCGAG 1311
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      392  LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln 411
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      1312  AGCGAGAACTTCGATAGCTGTCATTCAAGACAGACTCAAGGCCCGACATAGCCACCTC 1371
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      412  SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu 431
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      1372  GCGCGTGAACACTCCGTCATAGATAACCTGCCGAGAGAGGTGGTTGCCAAATTCATATGCG 1431
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      432  AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn 451
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      1432  CTTCAAGAGGAGCAGCAGGAGGAGGCTTAAGAACACACACCCCTTCAAGTTCTTCGTTCCA 1491
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      452  LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 471
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY      1492  CCGTCTCAGCAGCTCTCCGAGGCGCTGTGGCT 1521
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      472  ProLysGluSerGlnArgValValAla 481
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

RESULT 2
S11002
glycinin G2 precursor - soybean
N:Alternate names: glycinin A2B1a
C:Species: Glycine max (soybean)
C:Date: 21-Nov-1993 #sequence revision 19-Jan-1996 #text_change 05-Oct-2004
C:Accession: S11002; S04604; A26990
R:Nielsen, N.C.; Dickinson, C.D.; Cho, T.J.; Thanh, V.H.; Scallion, B.J.; Fischer, R.L.;
Plant Cell 1, 313-328, 1989
A:Title: Characterization of the glycinin gene family in soybean.
A:Reference number: S10851; MUID:92393391; PMID:2485233
A:Accession: S11002
A:Molecule type: DNA
A:Residues: 1-485 <NIE>
A:Cross-references: UNIPROT:P04405; UNIPARC:UPI000012B557
A:Experimental source: variety Dare
R:Thanh, V.H.; Tumer, N.E.; Nielsen, N.C.
Nucleic Acids Res. 17, 4387, 1989
A:Title: The glycinin Gy(2) gene from soybean.
A:Reference number: S04604; MUID:89296499; PMID:2740230
A:Accession: S04604
A:Molecule type: DNA
A:Residues: 1-485 <THA>
A:Cross-references: UNIPARC:UPI000012B557; EMBL:X15122; NID:g18636; PIDN:CAA33216.1; PI:
R:Fukazawa, C.; Momma, T.; Higuchi, W.; Udaka, K.
Nucleic Acids Res. 15, 8117, 1987
A:Title: Complete nucleotide sequence of the gene encoding a glycinin A2B1a subunit pre
A:Reference number: A26990; MUID:88040439; PMID:3671077
A:Accession: A26990
A:Molecule type: DNA
A:Residues: 1-485 <PUK>
A:Cross-references: UNIPARC:UPI000012B557; GB:X02806
C:Genetics:
A:Gene: GY2
A:Introns: 93/1; 177/3; 356/3
C:Keywords: storage protein
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-485/Product: glycinin G2 #status predicted <MAT>

Alignment Scores:
Pred. No.: 2.05e-103 Length: 485
Score: 1499.50 Matches: 297
Percent Similarity: 72.0% Conservative: 71
Best Local Similarity: 58.1% Mismatches: 94
Query Match: 54.0% Indels: 49
DB: 2 Gaps: 6

US-10-728-323-3 (1-1524) x S11002 (1-485)
QY      1  CGCGACCAACCGAGGAGAAACGCTGCCAGTTCACGCGCTCAATGCGCAGACCTGAC 60

```


Alignment Scores:				
Pred. No.:	3,74e-103	Length:	495	
Score:	1496.00	Matches:	292	
Percent Similarity:	70.9%	Conservative:	73	
Best Local Similarity:	56.7%	Mismatches:	100	
Query Match:	53.8%	Indels:	50	
DB:	2	Gaps:	6	
US-10-728-323-3 (1-1524) x S10851 (1-495)				
QY	1	CGGCACAAACCGGAGGAGAACGGCTGCCAGTTCCAGCGCTCAATGCGCAGAGACCTGAC	60	
DB	23	ArgGluGlnProGlnGlnAnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp	42	
QY	61	AATCGCATTGAATCAGAGCGCGTTACATTGAGACTTGGAACCCCAACACGAGGATTC	120	
DB	43	AsnArgIleGluSerGluGlyGlyLeuIleGluThrTriPAsnProAsnAsnLysProPhe	62	
QY	121	GAATGCGCGGCGTCCCTCTCTCGCTTAGTCTCCGCGGCAACGCCCTTCGTAGGCTT	180	
DB	63	GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro	82	
QY	181	TTCTACTCCAAATGCTCCAGGAGATCTTCATCCAGCAAGGAGGGGATACTTTGGGTG	240	
DB	83	SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyLysIlePheGlyMec	102	
QY	241	ATATTCCTCGTTGTCTTAGACACTATGAAGAGCCTCACACAAGGTCGTGATCTCAG	300	
DB	103	IleTyrProGlyCysProSerThrPheGluGluProGlnGlnProGlnArgGlyGln	122	
QY	301	TCCAAAGACCACCAAGACGTCTCCAAGGAGAGACCAAGCCAACAGCAACAGATAGT	360	
DB	123	SerSerArgPro-----GlnAspArg	129	
QY	361	CACCAGAAGTCACCGTTTCGATGAGGTGATCTCATTCGAGTCCACCGGTGTGCT	420	
DB	130	HisGlnLysIleTyrAsnPheArgGluGlyAspLeuIleAlaValProThrGlyValAla	149	
QY	421	TTCTGCTCTACACGACACGACACTGATGTGTGTCTCTTCTTACTGACACCAAC	480	
DB	150	TrpTrpMetTyrAsnAsnGluAspThrProValAlaValSerIleIleAspThrAsn	169	
QY	481	AACAACGACACACAGTTCATGATTCCTCCAGAGATTCATTTGGCTGGGACACGGAG	540	
DB	170	SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu	189	
QY	541	CAAGAGTTCCTTAAGGTACAGCAACAAAGCAGACAAAGCAGACGAAAGACTTACCATAT	600	
DB	190	GlnGluPheLeuLysTyrGlnGlnGluGln-----	199	
QY	601	AGCCCATACAGCCCGCAAAAGTCAGCCTAGACAAAGAGAGCGTGAAATTTAGCCCTCAGGA	660	
DB	200	-----GlyGly	201	
QY	661	CACACAGCGCAGAGAACGACGAGGACACAGAACAGAACAAACGAGTGAACATCTTC	720	
DB	202	HisGlnSerGlnLysGlyLysHisGlnGlnGluGluAsnGluGlyGlySerIleLeu	221	
QY	721	AGCGGTTTCACGCGGAGTTCCTGGAAACAGCCTTCAGGTTCACGACACAGATAGTG	780	
DB	222	SerGlyPheThrLeuGluPheLeuGluHisAlaPheSerVal---AspLysGlnIleAla	240	
QY	781	CAAAACCTAAGAGCGGACCGGAGGTGAAGAGAGAGGAGCCATTGTGACAGTGAAGGGA	840	
DB	241	LysAsnLeuGlnGlyGluAsnGluGlyGluAspLysGlyAlaIleValThrValLysGly	260	
QY	841	GGCTCAGATCTTGAGCCCA-----GATAGAAGAGAGCTCCGACGAGAGAGAG	891	
DB	261	GlyLeuSerValIleLysProProThrAspGluGlnGlnArgProGlnGluGluGlu	280	
QY	892	GAATACGATCAAGATGAATGAATAC-----GATGAAGAG---GATAGAAGG	936	

DB	281	GluGluGluGluAspGluLysProGlnCysLysGlyLysAspLysHisCysGlnArgPro	300	
QY	937	CGTGCAGGGGAAGACAGAGCAGGGGAATGTATTGAAGACAGATCTGCACCCGCAAGT	996	
DB	301	ArgGlySerGlnSerLysSerArgAsnGlyIleAspGluThrIleCysThrMetArg	320	
QY	997	GCTAAAAAGACATTGGTAGAAAACAGATCCCTCGATCATCTACAACCCCTCAAGCTGGTTCA	1056	
DB	321	LeuArgHisAsnIleGlyGlnThrSerSerProAspIleTyrAsnProGlnAlaGlySer	340	
QY	1057	CTCAAAACTGCCAACGATCTCAACCTTCTTAATACTTAGTGGCTTGGACTAGTGTGAA	1116	
DB	341	ValThrThrAlaThrSerLeuAspPheProAlaLeuSerTriPLeuArgLeuSerAlaGlu	360	
QY	1117	TATGMAATCTCTACAGGAATGCAATTTTTCGCTCACTACAACACCAACGACACACAGC	1176	
DB	361	PheGlySerLeuArgLysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSer	380	
QY	1177	ATCATATATTCGATTGAGGGGACGGCTCACGTGCAAGTCGTGGACAGCAACGCAACAGA	1236	
DB	381	IleIleTyrAlaLeuAsnGlyArgAlaLeuIleGlnValValAsnCysAsnGlyGluArg	400	
QY	1237	GTGTACGAGAGAGCTTCAAGAGGTCACGTCTTGTGTCGCCACAGAACTTCGCCGTC	1296	
DB	401	ValPheAspGlyGluLeuGlnGlyArgValLeuIleValProGlnAsnPheValVal	420	
QY	1297	GCTGMAATCCGACAGCGAGAACTTCGAATAGTGGCATTCAGACACACTCAAGCAGCCC	1356	
DB	421	AlaAlaArgSerGlnSerAspAsnPheGluTyrValSerPheLysThrAsnAspThrPro	440	
QY	1357	AGCATAGCCAACTCGCGGTGAAACTCCGTCAATAGATAAACCCTGCGGAGGAGTGGTT	1416	
DB	441	MetIleGlyThrLeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIle	460	
QY	1417	GCAATATCATATGGCTCCAAAGGAGGAGCAGCAAGCACTTAAGAACAAACCCCTTC	1476	
DB	461	GlnHisThrPheAsnLeuLysSerGlnGlnAlaArgGlnIleLysAsnAsnAsnProPhe	480	
QY	1477	AGTTCTTCCTCCACCGCTCAGCAGTCTCCGAGGGCTGTGGCT	1521	
DB	481	LysPheLeuValProProGlnGluSerGlnLysArgAlaValAla	495	
RESULT 4				
FWSYGI				
glycinin chain A2B1a precursor - soybean				
N;Alternate names: 11S globulin				
N;Contains: glycinin chain A2; glycinin chain B1a				
C;Species: Glycine max (soybean)				
C;Date: 31-Mar-1988 #sequence revision 31-Mar-1988 #text change 05-Oct-2004				
C;Accession: A91341; A92454; B92454; A90024; A90242; S10503; S74123; A05082; A05164; A0				
R;Mamma, T.; Negro, T.; Uda, K.; Fukazawa, C.				
FEBS Lett. 188, 117-122, 1985				
A;Title: A complete cDNA coding for the sequence of glycinin A2B1a subunit precursor.				
A;Reference number: A91341				
A;Accession: A91341				
A;Molecule type: mRNA				
A;Residues: 1-485 <MOM>				
A;Cross-references: UNIPROT:P04405; UNIPARC:UPI0000016DD98				
A;Experimental source: strain Bonimori				
A;Note: the source of this mRNA was cotyledon tissue taken from seeds at the middle sta				
R;Marco, Y.A.; Thanh, V.H.; Tumer, N.E.; Scallan, B.J.; Nielsen, N.C.				
J. Biol. Chem. 259, 13436-13441, 1984				
A;Title: Cloning and structural analysis of DNA encoding an A2B1a subunit of glycinin.				
A;Reference number: A92454; MUID:85030472; PMID:6092376				
A;Molecule type: mRNA				
A;Residues: 262-446 <MA1>				
A;Cross-references: UNIPARC:UPI000001745FF				
A;Accession: B92454				
A;Molecule type: DNA				
A;Residues: 318-485 <MA2>				
A;Cross-references: UNIPARC:UPI000001745FF				
A;Experimental source: strain CX635-1-1-1				

R:Utsumi, S.; Kim, C.S.; Kohno, M.; Kito, M.
 Agric. Biol. Chem. 51, 3267-3273, 1987
 A:Title: Polymorphism and expression of cDNAs encoding glycine subunits.
 A:Reference number: A90024
 A:Accession: A90024
 A:Molecule type: mRNA
 A:Residues: 1-38, 'D', '40-485 <UTS>
 A:Cross-references: UNIPARC:UPI000012B557
 A:Experimental source: strain Shiotsurunoko
 R:Staswick, P.E.; Hermodson, M.A.; Nielsen, N.C.
 J. Biol. Chem. 259, 13424-13430, 1984
 A:Title: The amino acid sequence of the A2B1a subunit of glycine.
 A:Reference number: A92452; MUID:85030470; PMID:6541652
 A:Accession: A92452
 A:Molecule type: protein
 A:Residues: 19-38, 'D', '40-60, 'S', '62-116, 'C', '118-192, 'E', '194-296, '301-342, 'S', '344-463, 'I', '465-485, 'D', '487-488, 'S', '490-491, 'D', '493-494, 'S', '496-497, 'D', '499-500, 'S', '502-503, 'D', '505-506, 'S', '508-509, 'D', '511-512, 'S', '514-515, 'D', '517-518, 'S', '520-521, 'D', '523-524, 'S', '526-527, 'D', '529-530, 'S', '532-533, 'D', '535-536, 'S', '538-539, 'D', '541-542, 'S', '544-545, 'D', '547-548, 'S', '550-551, 'D', '553-554, 'S', '556-557, 'D', '559-560, 'S', '562-563, 'D', '565-566, 'S', '568-569, 'D', '571-572, 'S', '574-575, 'D', '577-578, 'S', '580-581, 'D', '583-584, 'S', '586-587, 'D', '589-590, 'S', '592-593, 'D', '595-596, 'S', '598-599, 'D', '601-602, 'S', '604-605, 'D', '607-608, 'S', '610-611, 'D', '613-614, 'S', '616-617, 'D', '619-620, 'S', '622-623, 'D', '625-626, 'S', '628-629, 'D', '631-632, 'S', '634-635, 'D', '637-638, 'S', '640-641, 'D', '643-644, 'S', '646-647, 'D', '649-650, 'S', '652-653, 'D', '655-656, 'S', '658-659, 'D', '661-662, 'S', '664-665, 'D', '667-668, 'S', '670-671, 'D', '673-674, 'S', '676-677, 'D', '679-680, 'S', '682-683, 'D', '685-686, 'S', '688-689, 'D', '691-692, 'S', '694-695, 'D', '697-698, 'S', '700-701, 'D', '703-704, 'S', '706-707, 'D', '709-710, 'S', '712-713, 'D', '715-716, 'S', '718-719, 'D', '721-722, 'S', '724-725, 'D', '727-728, 'S', '730-731, 'D', '733-734, 'S', '736-737, 'D', '739-740, 'S', '742-743, 'D', '745-746, 'S', '748-749, 'D', '751-752, 'S', '754-755, 'D', '757-758, 'S', '760-761, 'D', '763-764, 'S', '766-767, 'D', '769-770, 'S', '772-773, 'D', '775-776, 'S', '778-779, 'D', '781-782, 'S', '784-785, 'D', '787-788, 'S', '790-791, 'D', '793-794, 'S', '796-797, 'D', '799-800, 'S', '802-803, 'D', '805-806, 'S', '808-809, 'D', '811-812, 'S', '814-815, 'D', '817-818, 'S', '820-821, 'D', '823-824, 'S', '826-827, 'D', '829-830, 'S', '832-833, 'D', '835-836, 'S', '838-839, 'D', '841-842, 'S', '844-845, 'D', '847-848, 'S', '850-851, 'D', '853-854, 'S', '856-857, 'D', '859-860, 'S', '862-863, 'D', '865-866, 'S', '868-869, 'D', '871-872, 'S', '874-875, 'D', '877-878, 'S', '880-881, 'D', '883-884, 'S', '886-887, 'D', '889-890, 'S', '892-893, 'D', '895-896, 'S', '898-899, 'D', '901-902, 'S', '904-905, 'D', '907-908, 'S', '910-911, 'D', '913-914, 'S', '916-917, 'D', '919-920, 'S', '922-923, 'D', '925-926, 'S', '928-929, 'D', '931-932, 'S', '934-935, 'D', '937-938, 'S', '940-941, 'D', '943-944, 'S', '946-947, 'D', '949-950, 'S', '952-953, 'D', '955-956, 'S', '958-959, 'D', '961-962, 'S', '964-965, 'D', '967-968, 'S', '970-971, 'D', '973-974, 'S', '976-977, 'D', '979-980, 'S', '982-983, 'D', '985-986, 'S', '988-989, 'D', '991-992, 'S', '994-995, 'D', '997-998, 'S', '1000-1001, 'D', '1003-1004, 'S', '1006-1007, 'D', '1009-1010, 'S', '1012-1013, 'D', '1015-1016, 'S', '1018-1019, 'D', '1021-1022, 'S', '1024-1025, 'D', '1027-1028, 'S', '1030-1031, 'D', '1033-1034, 'S', '1036-1037, 'D', '1039-1040, 'S', '1042-1043, 'D', '1045-1046, 'S', '1048-1049, 'D', '1051-1052, 'S', '1054-1055, 'D', '1057-1058, 'S', '1060-1061, 'D', '1063-1064, 'S', '1066-1067, 'D', '1069-1070, 'S', '1072-1073, 'D', '1075-1076, 'S', '1078-1079, 'D', '1081-1082, 'S', '1084-1085, 'D', '1087-1088, 'S', '1090-1091, 'D', '1093-1094, 'S', '1096-1097, 'D', '1099-1100, 'S', '1102-1103, 'D', '1105-1106, 'S', '1108-1109, 'D', '1111-1112, 'S', '1114-1115, 'D', '1117-1118, 'S', '1120-1121, 'D', '1123-1124, 'S', '1126-1127, 'D', '1129-1130, 'S', '1132-1133, 'D', '1135-1136, 'S', '1138-1139, 'D', '1141-1142, 'S', '1144-1145, 'D', '1147-1148, 'S', '1150-1151, 'D', '1153-1154, 'S', '1156-1157, 'D', '1159-1160, 'S', '1162-1163, 'D', '1165-1166, 'S', '1168-1169, 'D', '1171-1172, 'S', '1174-1175, 'D', '1177-1178, 'S', '1180-1181, 'D', '1183-1184, 'S', '1186-1187, 'D', '1189-1190, 'S', '1192-1193, 'D', '1195-1196, 'S', '1198-1199, 'D', '1201-1202, 'S', '1204-1205, 'D', '1207-1208, 'S', '1210-1211, 'D', '1213-1214, 'S', '1216-1217, 'D', '1219-1220, 'S', '1222-1223, 'D', '1225-1226, 'S', '1228-1229, 'D', '1231-1232, 'S', '1234-1235, 'D', '1237-1238, 'S', '1240-1241, 'D', '1243-1244, 'S', '1246-1247, 'D', '1249-1250, 'S', '1252-1253, 'D', '1255-1256, 'S', '1258-1259, 'D', '1261-1262, 'S', '1264-1265, 'D', '1267-1268, 'S', '1270-1271, 'D', '1273-1274, 'S', '1276-1277, 'D', '1279-1280, 'S', '1282-1283, 'D', '1285-1286, 'S', '1288-1289, 'D', '1291-1292, 'S', '1294-1295, 'D', '1297-1298, 'S', '1300-1301, 'D', '1303-1304, 'S', '1306-1307, 'D', '1309-1310, 'S', '1312-1313, 'D', '1315-1316, 'S', '1318-1319, 'D', '1321-1322, 'S', '1324-1325, 'D', '1327-1328, 'S', '1330-1331, 'D', '1333-1334, 'S', '1336-1337, 'D', '1339-1340, 'S', '1342-1343, 'D', '1345-1346, 'S', '1348-1349, 'D', '1351-1352, 'S', '1354-1355, 'D', '1357-1358, 'S', '1360-1361, 'D', '1363-1364, 'S', '1366-1367, 'D', '1369-1370, 'S', '1372-1373, 'D', '1375-1376, 'S', '1378-1379, 'D', '1381-1382, 'S', '1384-1385, 'D', '1387-1388, 'S', '1390-1391, 'D', '1393-1394, 'S', '1396-1397, 'D', '1399-1400, 'S', '1402-1403, 'D', '1405-1406, 'S', '1408-1409, 'D', '1411-1412, 'S', '1414-1415, 'D', '1417-1418, 'S', '1420-1421, 'D', '1423-1424, 'S', '1426-1427, 'D', '1429-1430, 'S', '1432-1433, 'D', '1435-1436, 'S', '1438-1439, 'D', '1441-1442, 'S', '1444-1445, 'D', '1447-1448, 'S', '1450-1451, 'D', '1453-1454, 'S', '1456-1457, 'D', '1459-1460, 'S', '1462-1463, 'D', '1465-1466, 'S', '1468-1469, 'D', '1471-1472, 'S', '1474-1475, 'D', '1477-1478, 'S', '1480-1481, 'D', '1483-1484, 'S', '1486-1487, 'D', '1489-1490, 'S', '1492-1493, 'D', '1495-1496, 'S', '1498-1499, 'D', '1501-1502, 'S', '1504-1505, 'D', '1507-1508, 'S', '1510-1511, 'D', '1513-1514, 'S', '1516-1517, 'D', '1519-1520, 'S', '1522-1523, 'D', '1525-1526, 'S', '1528-1529, 'D', '1531-1532, 'S', '1534-1535, 'D', '1537-1538, 'S', '1540-1541, 'D', '1543-1544, 'S', '1546-1547, 'D', '1549-1550, 'S', '1552-1553, 'D', '1555-1556, 'S', '1558-1559, 'D', '1561-1562, 'S', '1564-1565, 'D', '1567-1568, 'S', '1570-1571, 'D', '1573-1574, 'S', '1576-1577, 'D', '1579-1580, 'S', '1582-1583, 'D', '1585-1586, 'S', '1588-1589, 'D', '1591-1592, 'S', '1594-1595, 'D', '1597-1598, 'S', '1600-1601, 'D', '1603-1604, 'S', '1606-1607, 'D', '1609-1610, 'S', '1612-1613, 'D', '1615-1616, 'S', '1618-1619, 'D', '1621-1622, 'S', '1624-1625, 'D', '1627-1628, 'S', '1630-1631, 'D', '1633-1634, 'S', '1636-1637, 'D', '1639-1640, 'S', '1642-1643, 'D', '1645-1646, 'S', '1648-1649, 'D', '1651-1652, 'S', '1654-1655, 'D', '1657-1658, 'S', '1660-1661, 'D', '1663-1664, 'S', '1666-1667, 'D', '1669-1670, 'S', '1672-1673, 'D', '1675-1676, 'S', '1678-1679, 'D', '1681-1682, 'S', '1684-1685, 'D', '1687-1688, 'S', '1690-1691, 'D', '1693-1694, 'S', '1696-1697, 'D', '1699-1700, 'S', '1702-1703, 'D', '1705-1706, 'S', '1708-1709, 'D', '1711-1712, 'S', '1714-1715, 'D', '1717-1718, 'S', '1720-1721, 'D', '1723-1724, 'S', '1726-1727, 'D', '1729-1730, 'S', '1732-1733, 'D', '1735-1736, 'S', '1738-1739, 'D', '1741-1742, 'S', '1744-1745, 'D', '1747-1748, 'S', '1750-1751, 'D', '1753-1754, 'S', '1756-1757, 'D', '1759-1760, 'S', '1762-1763, 'D', '1765-1766, 'S', '1768-1769, 'D', '1771-1772, 'S', '1774-1775, 'D', '1777-1778, 'S', '1780-1781, 'D', '1783-1784, 'S', '1786-1787, 'D', '1789-1790, 'S', '1792-1793, 'D', '1795-1796, 'S', '1798-1799, 'D', '1801-1802, 'S', '1804-1805, 'D', '1807-1808, 'S', '1810-1811, 'D', '1813-1814, 'S', '1816-1817, 'D', '1819-1820, 'S', '1822-1823, 'D', '1825-1826, 'S', '1828-1829, 'D', '1831-1832, 'S', '1834-1835, 'D', '1837-1838, 'S', '1840-1841, 'D', '1843-1844, 'S', '1846-1847, 'D', '1849-1850, 'S', '1852-1853, 'D', '1855-1856, 'S', '1858-1859, 'D', '1861-1862, 'S', '1864-1865, 'D', '1867-1868, 'S', '1870-1871, 'D', '1873-1874, 'S', '1876-1877, 'D', '1879-1880, 'S', '1882-1883, 'D', '1885-1886, 'S', '1888-1889, 'D', '1891-1892, 'S', '1894-1895, 'D', '1897-1898, 'S', '1900-1901, 'D', '1903-1904, 'S', '1906-1907, 'D', '1909-1910, 'S', '1912-1913, 'D', '1915-1916, 'S', '1918-1919, 'D', '1921-1922, 'S', '1924-1925, 'D', '1927-1928, 'S', '1930-1931, 'D', '1933-1934, 'S', '1936-1937, 'D', '1939-1940, 'S', '1942-1943, 'D', '1945-1946, 'S', '1948-1949, 'D', '1951-1952, 'S', '1954-1955, 'D', '1957-1958, 'S', '1960-1961, 'D', '1963-1964, 'S', '1966-1967, 'D', '1969-1970, 'S', '1972-1973, 'D', '1975-1976, 'S', '1978-1979, 'D', '1981-1982, 'S', '1984-1985, 'D', '1987-1988, 'S', '1990-1991, 'D', '1993-1994, 'S', '1996-1997, 'D', '1999-2000, 'S', '2002-2003, 'D', '2005-2006, 'S', '2008-2009, 'D', '2011-2012, 'S', '2014-2015, 'D', '2017-2018, 'S', '2020-2021, 'D', '2023-2024, 'S', '2026-2027, 'D', '2029-2030, 'S', '2032-2033, 'D', '2035-2036, 'S', '2038-2039, 'D', '2041-2042, 'S', '2044-2045, 'D', '2047-2048, 'S', '2050-2051, 'D', '2053-2054, 'S', '2056-2057, 'D', '2059-2060, 'S', '2062-2063, 'D', '2065-2066, 'S', '2068-2069, 'D', '2071-2072, 'S', '2074-2075, 'D', '2077-2078, 'S', '2080-2081, 'D', '2083-2084, 'S', '2086-2087, 'D', '2089-2090, 'S', '2092-2093, 'D', '2095-2096, 'S', '2098-2099, 'D', '2101-2102, 'S', '2104-2105, 'D', '2107-2108, 'S', '2110-2111, 'D', '2113-2114, 'S', '2116-2117, 'D', '2119-2120, 'S', '2122-2123, 'D', '2125-2126, 'S', '2128-2129, 'D', '2131-2132, 'S', '2134-2135, 'D', '2137-2138, 'S', '2140-2141, 'D', '2143-2144, 'S', '2146-2147, 'D', '2149-2150, 'S', '2152-2153, 'D', '2155-2156, 'S', '2158-2159, 'D', '2161-2162, 'S', '2164-2165, 'D', '2167-2168, 'S', '2170-2171, 'D', '2173-2174, 'S', '2176-2177, 'D', '2179-2180, 'S', '2182-2183, 'D', '2185-2186, 'S', '2188-2189, 'D', '2191-2192, 'S', '2194-2195, 'D', '2197-2198, 'S', '2200-2201, 'D', '2203-2204, 'S', '2206-2207, 'D', '2209-2210, 'S', '2212-2213, 'D', '2215-2216, 'S', '2218-2219, 'D', '2221-2222, 'S', '2224-2225, 'D', '2227-2228, 'S', '2230-2231, 'D', '2233-2234, 'S', '2236-2237, 'D', '2239-2240, 'S', '2242-2243, 'D', '2245-2246, 'S', '2248-2249, 'D', '2251-2252, 'S', '2254-2255, 'D', '2257-2258, 'S', '2260-2261, 'D', '2263-2264, 'S', '2266-2267, 'D', '2269-2270, 'S', '2272-2273, 'D', '2275-2276, 'S', '2278-2279, 'D', '2281-2282, 'S', '2284-2285, 'D', '2287-2288, 'S', '2290-2291, 'D', '2293-2294, 'S', '2296-2297, 'D', '2299-2300, 'S', '2302-2303, 'D', '2305-2306, 'S', '2308-2309, 'D', '2311-2312, 'S', '2314-2315, 'D', '2317-2318, 'S', '2320-2321, 'D', '2323-2324, 'S', '2326-2327, 'D', '2329-2330, 'S', '2332-2333, 'D', '2335-2336, 'S', '2338-2339, 'D', '2341-2342, 'S', '2344-2345, 'D', '2347-2348, 'S', '2350-2351, 'D', '2353-2354, 'S', '2356-2357, 'D', '2359-2360, 'S', '2362-2363, 'D', '2365-2366, 'S', '2368-2369, 'D', '2371-2372, 'S', '2374-2375, 'D', '2377-2378, 'S', '2380-2381, 'D', '2383-2384, 'S', '2386-2387, 'D', '2389-2390, 'S', '2392-2393, 'D', '2395-2396, 'S', '2398-2399, 'D', '2401-2402, 'S', '2404-2405, 'D', '2407-2408, 'S', '2410-2411, 'D', '2413-2414, 'S', '2416-2417, 'D', '2419-2420, 'S', '2422-2423, 'D', '2425-2426, 'S', '2428-2429, 'D', '2431-2432, 'S', '2434-2435, 'D', '2437-2438, 'S', '2440-2441, 'D', '2443-2444, 'S', '2446-2447, 'D', '2449-2450, 'S', '2452-2453, 'D', '2455-2456, 'S', '2458-2459, 'D', '2461-2462, 'S', '2464-2465, 'D', '2467-2468, 'S', '2470-2471, 'D', '2473-2474, 'S', '2476-2477, 'D', '2479-2480, 'S', '2482-2483, 'D', '2485-2486, 'S', '2488-2489, 'D', '2491-2492, 'S', '2494-2495, 'D', '2497-2498, 'S', '2500-2501, 'D', '2503-2504, 'S', '2506-2507, 'D', '2509-2510, 'S', '2512-2513, 'D', '2515-2516, 'S', '2518-2519, 'D', '2521-2522, 'S', '2524-2525, 'D', '2527-2528, 'S', '2530-2531, 'D', '2533-2534, 'S', '2536-2537, 'D', '2539-2540, 'S', '2542-2543, 'D', '2545-2546, 'S', '2548-2549, 'D', '2551-2552, 'S', '2554-2555, 'D', '2557-2558, 'S', '2560-2561, 'D', '2563-2564, 'S', '2566-2567, 'D', '2569-2570, 'S', '2572-2573, 'D', '2575-2576, 'S', '2578-2579, 'D', '2581-2582, 'S', '2584-2585, 'D', '2587-2588, 'S', '2590-2591, 'D', '2593-2594, 'S', '2596-2597, 'D', '2599-2600, 'S', '2602-2603, 'D', '2605-2606, 'S', '2608-2609, 'D', '2611-2612, 'S', '2614-2615, 'D', '2617-2618, 'S', '2620-2621, 'D', '2623-2624, 'S', '2626-2627, 'D', '2629-2630, 'S', '2632-2633, 'D', '2635-2636, 'S', '2638-2639, 'D', '2641-2642, 'S', '2644-2645, 'D', '2647-2648, 'S', '2650-2651, 'D', '2653-2654, 'S', '2656-2657, 'D', '2659-2660, 'S', '2662-2663, 'D', '2665-2666, 'S', '2668-2669, 'D', '2671-2672, 'S', '2674-2675, 'D', '2677-2678, 'S', '2680-2681, 'D', '2683-2684, 'S', '2686-2687, 'D', '2689-2690, 'S', '2692-2693, 'D', '2695-2696, 'S', '2698-2699, 'D', '2701-2702, 'S', '2704-2705, 'D', '2707-2708, 'S', '2710-2711, 'D', '2713-2714, 'S', '2716-2717, 'D', '2719-2720, 'S', '2722-2723, 'D', '2725-2726, 'S', '2728-2729, 'D', '2731-2732, 'S', '2734-2735, 'D', '2737-2738, 'S', '2740-2741, 'D', '2743-2744, 'S', '2746-2747, 'D', '2749-2750, 'S', '2752-2753, 'D', '2755-2756, 'S', '2758-2759, 'D', '2761-2762, 'S', '2764-2765, 'D', '2767-2768, 'S', '2770-2771, 'D', '2773-2774, 'S', '2776-2777, 'D', '2779-2780, 'S', '2782-2783, 'D', '2785-2786, 'S', '2788-2789, 'D', '2791-2792, 'S', '2794-2795, 'D', '2797-2798, 'S', '2800-2801, 'D', '2803-2804, 'S', '2806-2807, 'D', '2809-2810, 'S', '2812-2813, 'D', '2815-2816, 'S', '2818-2819, 'D', '2821-2822, 'S', '2824-2825, 'D', '2827-2828, 'S', '2830-2831, 'D', '2833-2834, 'S', '2836-2837, 'D', '2839-2840, 'S', '2842-2843, 'D', '2845-2846, 'S', '2848-2849, 'D', '2851-2852, 'S', '2854-2855, 'D', '2857-2858, 'S', '2860-2861, 'D', '2863-2864, 'S', '2866-2867, 'D', '2869-2870, 'S', '2872-2873, 'D', '2875-2876, 'S', '2878-2879, 'D', '2881-2882, 'S', '2884-2885, 'D', '2887-2888, 'S', '2890-2891, 'D', '2893-2894, 'S', '2896-2897, 'D', '2899-2900, 'S', '2902-2903, 'D', '2905-2906, 'S', '2908-2909, 'D', '2911-2912, 'S', '2914-2915, 'D', '2917-2918, 'S', '2920-2921, 'D', '2923-2924, 'S', '2926-2927, 'D', '2929-2930, 'S', '2932-2933, 'D', '2935-2936, 'S', '2938-2939, 'D', '2941-2942, 'S', '2944-2945, 'D', '2947-2948, 'S', '2950-2951, 'D', '2953-2954, 'S', '2956-2957, 'D', '2959-2960, 'S', '2962-2963, 'D', '2965-2966, 'S', '2968-2969, 'D', '2971-2972, 'S', '2974-2975, 'D', '2977-2978, 'S', '2980-2981, 'D', '2983-2984, 'S', '2986-2987, 'D', '2989-2990, 'S', '2992-2993, 'D', '2995-2996, 'S', '2998-2999, 'D', '3001-3002, 'S', '3004-3005, 'D', '3007-3008, 'S', '3010-3011, 'D', '3013-3014, 'S', '3016-3017, 'D', '3019-3020, 'S', '3022-3023, 'D', '3025-3026, 'S', '3028-3029, 'D', '3031-3032, 'S', '3034-3035, 'D', '3037-3038, 'S', '3040-3041, 'D', '3043-3044, 'S', '3046-3047, 'D', '3049-3050, 'S', '3052-3053, 'D', '3055-3056, 'S', '3058-3059, 'D', '3061-3062, 'S', '3064-3065, 'D', '3067-3068, 'S', '3070-3071, 'D', '3073-3074, 'S', '3076-3077, 'D', '3079-3080, 'S', '3082-3083, 'D', '3085-3086, 'S',

```

Db      375  LeuAsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheaspGly 394
QY      1249  GAGCTTCAAGAGGTCACGTGCTTGTGGTGGCCACAGAACTTCGCGCTCGTGGAAAGTCC 1308
Db      395  GluLeuGlnGluGlyValLeuIleValProGlnAsnPheAlaValAlaIalysSer 414
QY      1309  CAGAGCGAAGCTTCAATACGTGGATTTCAGACAGACTCAAGGCCAGCATAGCCAAC 1368
Db      415  GlnSerAspAsnPhGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsn 434
QY      1369  CTCGCGGTGAACCTCCGCTCATAGATAACCTCCGCGAGGAGGTGTTGCAAAATTCATAT 1428
Db      435  LeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnHisThrPhe 454
QY      1429  GGCCTTCAAGGAGCAGGACGAGCTTAAAGAACAAACACCCCTTCAAGTTCTTCGTT 1488
Db      455  AsnLeuLysSerGlnGlnAlaArgGlnValLysAsnAsnAsnPropheSerPheLeuVal 474
QY      1489  CCACCGTCTCAGCAGTCTCCGAGGCGTGGCT 1521
Db      475  ProProGlnGluSerGlnArgArgAlaValAla 485

RESULT 5
FWSYG2
glycinin chain AlaBx precursor - soybean
N;Alternate names: 11S globulin; glycinin AlaB1b
C;Species: Glycine max (soybean)
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 05-Oct-2004
R;Accession: A23497; S10502
R;Negoro, T.; Momma, T.; Fukazawa, C.
Nucleic Acids Res. 13, 6719-6731, 1985
A;Title: A cDNA clone encoding a glycinin A-1a subunit precursor of soybean.
A;Reference number: A23497; MUID:86041867; PMID:2997720
A;Accession: A23497
A;Molecule type: mRNA
A;Cross-references: UNIPROT:P04776; UNIPARC:UPI0000042425; GB:X02985; NID:g18614; PIDN:C
A;Experimental source: cv. Boninmori
A;Note: The authors translated the codon AAC for residue 449 as Lys
R;Kitamura, Y.; Arahira, M.; Itoh, Y.; Fukazawa, C.
Nucleic Acids Res. 18, 4245, 1990
A;Title: The complete nucleotide sequence of soybean glycinin A2B1a gene spanning to and
A;Reference number: S10502; MUID:90332420; PMID:2377465
A;Accession: S10502
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 481-495 <KIT>
A;Cross-references: UNIPARC:UPI00000067E; EMBL:X53404; NID:g18522; PIDN:CAA37479.1; PID
C;Comment: The source of this protein was cotyledon tissue taken 38 days after flowering
C;Keywords: seed; storage protein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-306/Product: glycinin chain Ala #status predicted <GLA>
F;311-490/Product: glycinin chain Bx #status predicted <GLB>
F;107-317/Disulfide bonds: #status predicted

Alignment Scores:
Pred. No.: 6.32e-101 Length: 495
Score: 1466.00 Matches: 288
Percent Similarity: 70.1% Conservative: 73
Best Local Similarity: 55.9% Mismatches: 104
Query Match: 52.8% Indels: 50
DB: 1 Gaps: 6

US-10-728-323-3 (1-1524) x FWSYG2 (1-495)

QY      1  CGCAGCAACCGGAGGACCGTCCAGTTCAGCGCTCAATGGCAGAGACCTGAC 60
Db      23  ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProGly 42
QY      61  AATCGCATTTGAATCAGAGGCGGTATCATTTGAGACTTGGAACTTGAACCCCAACACACGAGGATTC 120

```

```

Db      43  AsnArgIleGluSerGluGlyGlyLeuIleGluThrTrpAsnProAsnAsnLysProPhe 62
QY      121  GAATGGCGGCGTGCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCCT 180
Db      63  GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 82
QY      181  TTCTACTCAAGCTCCCGCAGAGATCTTCATCCAGCAGGAGGCGATACTTTGGTTG 240
Db      83  SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyLysGlyIlePheGlyMet 102
QY      241  ATATTCCCTGGTGTCTTAGACACATATGAAGAGCCTCACACAAAGTCTCGATCTCAG 300
Db      103  IleTyrProGlyCysSerSerThrPheGluGluProGlnGlnProGlnGlnArgGlyGln 122
QY      301  TCCAAAAGACCAACAGAGCTCTCCAAGGAGAAGCAACCAAGCCACAGCAACAGATAGT 360
Db      123  SerSerArgPro-----GlnAspArg 129
QY      361  CACCAAGAGGTGCACCGTTTCGATGAGGTGATCTCATTGCAGTTCCACACCGGTGTGCT 420
Db      130  HisGlnLysIleTyrAsnSerArgGluGlyAspLeuIleAlaValProThrGlyValAla 149
QY      421  TTCTGCTCTACACGACCAACGACACTGATGTTGTTGCTGTTCTTCTTACTGACCAAC 480
Db      150  TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleIleAspThrAsn 169
QY      481  AACACGACCAACAGCTTGATCAGTCCCGCAGAGATTCAATTGGCTGGGACACGGAG 540
Db      170  SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 189
QY      541  CAAGAGTCTTAAAGGTACCAGCAACAAAGCAGACAAAGCAGACAGCAAGAAAGCTTACCATAT 600
Db      190  GlnGluPheLeuLysTyrGlnGlnGln----- 199
QY      601  AGCCCATACAGCCCGCAAAAGTCAGCCTAGACAAAGAGAGCGTGAATTGATCCCTCGAGGA 660
Db      200  -----GlyGly 201
QY      661  CAGCACAGCGGAGAGAACGACGAGGACAAGAGAAGAAAACGAAGTGGAAACATCTTC 720
Db      202  HisGlnSerGlnLysGlyLysHisGlnGlnGluGluAsnGluGlySerIleLeu 221
QY      721  AGCGGCTTCAACGCGGAGTCTCTGGAAACAAGCTCTCCAGTTCAGCAGACAGATAGTG 780
Db      222  SerGlyPheThrLeuGluPheLeuGluHisAlaPheSerVal---AspLysGlnIleAla 240
QY      781  CAAACCTAAGAGCGGAGACCGAGAGTGAAGAGAGGAGGCCATTTGTAGTGGGGA 840
Db      241  LysAsnLeuGlnGlyGluAsnGluGlyGluAspLysGlyAlaIleValThrValLysGly 260
QY      841  GGCCTCAGAACTCTTGAGCCCA-----GATAGAAAGAGAGCTCCGACGAGAGAGAG 891
Db      261  GlyLeuSerValIleLysProThrAspGluGlnGlnGlnArgProGlnGluGlu 280
QY      892  GAATACGATCAAGATGAATATGAATAC-----GATGAAGAG---GATAGAAGG 936
Db      281  GluGluGluLysGluLysProGlnCysLysGlyLysAspLysHisCysGlnArgPro 300
QY      937  CGTGGCAGGGAGCAGAGGCGAGGGAATGTGATTGAAGAGACGATCTGCACCGCAAGT 996
Db      301  ArgGlySerGlnSerLysSerArgArgAsnGlyIleAspGluThrIleCysThrMetArg 320
QY      997  GCTAAAAGAACNTTGGTAGAAACAGATCCCTTGACATCTACACCTCAAGCTGTTCA 1056
Db      321  LeuArgHisAsnIleGlyGlnThrSerSerProAspIleTyrAsnProGlnAlaGlySer 340
QY      1057  CTCAAACTGCCAACCGATCTCAACCTTCTAACTTAGTGGCTTGACCTAGTGTGAA 1116
Db      341  ValThrThrAlaThrSerLeuAspPheProAlaLeuSerTrpLeuArgLeuSerAlaGly 360
QY      1117  TATGGAATCTCTACAGGAATGCAATTGTTGTCGCTCCTACCAACCAACGACACAGC 1176
Db      361  PheGlySerLeuArgLysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSer 380

```


[illegible]

RESULT 7

S08237
legumin A2 precursor - garden pea
N:Alternate names: 11S seed storage protein
C:Species: Pisum sativum (garden pea)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 05-Oct-2004
C:Accession: S08237
R:Rerie, W.G.; Whitecross, M.I.; Higgins, T.J.V.
Nucleic Acids Res. 18, 655, 1990
A:Title: Nucleotide sequence of an A-type legumin gene from pea.
A:Reference number: S08237; MUID:90174993; PMID:2308850
A:Accession: S08237
A:Molecule type: DNA
A:Residues: 1-520 <RER>
A:Cross-references: UNIPROT:P15838; UNIPARC:UPI0000012E424; EMBL:X17193; NID:920773; PIDN:
C:Genetics:
A:Gene: legA2
A:Introns: 97/1; 180/3; 391/3
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-335/Product: legumin A2 alpha chain #status predicted <ACH>
F:336-520/Product: legumin A2 beta chain #status predicted <BCH>

Alignment Scores:		
Pred. No.:	4.17e-100	Length: 520
Score:	1455.00	Matches: 295
Percent Similarity:	65.5%	Conservative: 64
Best Local Similarity:	53.8%	Mismatches: 97
Query Match:	52.4%	Indels: 92
DB:	2	Gaps: 8
US-10-728-323-3 (1-1524)	x	S08237 (1-520)

[illegible]

892	QY	GAATACGATCAAGATGAATATGATTAATACGATCAAGAGGATAGAGCGCTGGC-----	942
292	Db	GlulLaspGluGluArgGlnProArgHisGlnArgArgArgGlyGluGluGlu 311	
943	QY	-----AGGGGAAGCAGAGGCAGGGGG-----	963
312	Db	GlulLaspLysLysGluArgGlyGlySerGlnLysGlyLysSerArgArgGlnGlyAsp 331	
964	QY	AATGGTATTGAAGAGACGATCTCCACCGCAAGTGCCTAAAGAACAATGTGTAGAAACAGA 1022	
332	Db	AsnGlyLeuGluGluThrValCysThrAlaLysLeuArgLeuAsnIleGlyProSerSer 351	
1024	QY	TCCCTCGACATCTACAAACCTCAAGCTGGTTCTACTCAAACTGCCAACGATCTCAACCTT 108	
352	Db	SerProAspIleTyrAsnProGluAlaGlyArgLysLysThrValThrSerLeuAspLeu 371	
1084	QY	CTAATCTACTAGGTGGCTTGGACCTAGTGTCTGAATATGTAATCTCTACAGGAATGCATTT 1143	
372	Db	ProValLeuArgTrpLeuLysLeuSerAlaGluHisGlySerLeuHisLysAsnAlaMet 391	
1144	QY	TTTGTCTGCTCACTACAAACACCAAGCACAGCATCATATATCGATTGAGGGGACGGCT 1203	
392	Db	PheValProHisTyrAsnLeuAsnAlaAsnSerIleTyrAlaLeuLysGlyArgAla 411	
1204	QY	CACGTGCAAGTCTGTGACAGACCAACGGCAACAGAGTGTACACAGGAGAGCTTCAAGAGGGT 1263	
412	Db	ArgLeuGlnValValAsnCysAsnGlyAsnThrValPheaspGlyGluLeuGluAlaGly 431	
1264	QY	CAGTGTCTGTGTGGCCACAGAATCTCGCGTCTGCTGGAAAGTCCCAAGACGGAGAACTTC 1322	
432	Db	ArgAlaLeuThrValProGlnAsnTyrAlaValAlaAlaLysSerLeuSerAspArgPhe 451	
1324	QY	GAATAGCTGCATTCAAGACAGACTCAAGGCCCAAGCATGCCAACCTTCGCGGTGAAAC 138	
452	Db	SerTyrValAlaPheLysThrAsnAspArgAlaGlyIleAlaArgLeuAlaGlyThrSer 471	
1384	QY	TCCGTTCATAGATAACCTGCCGGAGGGGTGGTTGCCAAATTCATATGCCTCCAAAGGGAG 1443	

```
Db 472 SerValIleAsnAsnLeuProLeuAspValAlaAlaThrPheAsnLeuGlnArgAsn 491
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1444 CAGCGAAGGAGGAGTAAAGAACAAACCCCTTCAAGTTCTTCCACCGTCTCAGCAG 1503
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 492 GluAlaArgGlnLeuLysSerAsnAsnProPheLysPheLeuValProAlaArgGluSer 511
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1504 TCTCCAGGCGCTGTGGCT 1521
|||||
Db 512 GluAsnArgAlaSerAla 517
|||||

RESULT 9
S49877
legumin A precursor - Vicia narbonensis
C:Species: Vicia narbonensis
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-Oct-2004
C:Accession: S49877
R:Nong, V.; Becker, C.; Muentz, K.
submitted to the EMBL Data Library, November 1994
A:Description: PCR cloning of legumin cDNA from Vicia narbonensis.
A:Reference number: S49877
A:Accession: S49877
A:Molecule type: mRNA
A:Residues: 1-482 <NON>
A:Cross-references: UNIPROT:Q41676; UNIPARC:UPI000000A6C09; EMBL:Z46803; NID:g600107; PII
A:Experimental source: tissue type cotyledon
C:Keywords: seed; storage protein
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-482/Product: legumin A #status predicted <MAT>

Alignment Scores:
Pred. No.: 2,3e-98 Length: 482
Score: 1431.50 Matches: 284
Percent Similarity: 68.2% Conservative: 66
Best Local Similarity: 55.4% Mismatches: 104
Query Match: 51.5% Indels: 59
DB: 2 Gaps: 7

US-10-728-323-3 (1-1524) x S49877 (1-482)
QY 1 CGGCAGCAACCGGAGAGAACGGGTGCCAGTTCAGCGCTCAATGGCAGACCTGAC 60
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 23 ArgGluGlnSerGlnGlnAsnGluCysGlnLeuGluArgLeuAspAlaLeuGluProAsp 42
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 61 AATCGCATTAATCAGAGGCGGTTCATTGAGACTTGGAACCCCAACCAACAGGAGTTC 120
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 43 AsnArgIleGluSerGluGlyGlyLeuIleGluThrTriPAsnProAsnAsnArgGlnPhe 62
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 121 GAATGCGCGCGCTCGCCCTCTCGCTTAGTCTCGCGCGCAACGCCCTTCGTAGGCCT 180
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 63 ArgCysAlaGlyValAlaLeuSerArgValThrLeuGlnArgAsnAlaLeuArgArgPro 82
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 181 TTCTACTCCATGCTCCCAAGAGATCTTCATCCAGCAAGGAAGGGGATACTTTGGGTG 240
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 83 TyrTyrSerAsnAlaProGlnGluIleTyrIleGlnGlnGlyAsnGlyTyrPheGlyVal 102
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 241 ATATTCCCTGGTGTCTTAGACACTATGAGAGCGCTCACACAAAGTCTGCTCATCTCAG 300
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 103 ValPheProGlyCysProGluThrPheGluGluPro-----GlnGluSer 117
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 301 TCCCAAGACCAACCAAGACCTCTCCAAGGAGAGAACCAAGCAACAGCAACAGATAGT 360
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 118 GluGlnArgGluArgArg------TyrArgAspSer 128
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 361 CACCAGAAGGTGCACCGCTTTTCGATGAGGTGATCTCATTCAGTTCCTCCACCGGTGTGCT 420
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 129 HisGlnLysValAsnArgPheArgGluGlyAspIleAlaValProThrGlyAsnVal 148
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 421 TTCTGGCTCTAACACGACACGACACTGATGTGTGCTGCTTCTCTTACTACACCAAC 480
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 149 LeuTriPMetTyrAsnAspGlnAspThrProValIleAlaIleSerLeuThrAspThrGly 168
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 481 AACACGACCAACCAAGCTTGATCATGTTCCCGAGGAGATTCAATTTGGCTGGGAACACGAG 540
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
```

RESULT 10
S11003

glycinin G3 precursor - soybean
 C:Species: Glycine max (soybean)
 C:Date: 21-Nov-1993 #sequence_revision 19-Jan-1996 #text_change 05-Oct-2004
 C:Accession: S11003
 R:Nielsen, N.C.; Dickinson, C.D.; Cho, T.J.; Thanh, V.H.; Scallion, B.J.; Fischer, R.L.;
 Plant Cell 1, 313-328, 1989
 A:Title: Characterization of the glycinin gene family in soybean.
 A:Reference number: S10851; MUID:92393391; PMID:2485233
 A:Accession: S11003
 A:Molecule type: DNA
 A:Residues: 1-484 <NIE>
 A:Cross-references: UNIPARC:UPI0000177DFE
 A:Experimental source: variety Dare
 C:Genetics:
 A:Gene: GY3
 C:Keywords: storage protein
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-484/Product: glycinin G3 #status predicted <NAT>

Alignment Scores:
 Pred. No.: 4.18e-98 Length: 484
 Score: 1428.00 Matches: 290
 Percent Similarity: 70.3% Conservative: 68
 Best Local Similarity: 57.0% Mismatches: 95
 Query Match: 51.4% Indels: 56
 DB: 2 Gaps: 9

US-10-728-323-3 (1-1524) x S11003 (1-484)

QY	1	CGGAGCAACCGAGAGAGAAACGGTGCCTCCAGCTTCCAGCGCTCAATGCGGAGACCTGAC	60
DB	23	ArgGluGlnProGlnAsnGluCysGlnLeuArgLeuAsnAlaLeuLysProAsp	42
QY	61	AATCGATTGAATCAGAGCGGCTTACATTGAGCTTGGAACCCCAACACACGAGGATC	120
DB	43	AsnArgileGluSerGluGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe	62
QY	121	GAATGCGCGGCTGCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGGCT	180
DB	63	GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro	82
QY	181	TTCTACTCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGATCTTTGGTTG	240
DB	83	SerTyrThrAsnAlaProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet	102
QY	241	ATATTCCTGTTGTCTAGACACTATGAGAGCCTCACACAGGTCGTGATCTCAG	300
DB	103	IlePheProGlyCysProSerThrPheGluGluProGlnGlnLysGly-----Gln	119
QY	301	TCCAAAGACCAACCAAGAGCTCTCCAAGGAGAAGACCAAGCCACAGCAACGAGATAGT	360
DB	120	SerSerArgPro-----GlnAspArg	126
QY	361	CACCAAGAGTGCACCGTTTCGATGAGGTGATCTCATTCGATTCCTCCACCGTGTGCT	420
DB	127	HisGlnLysIleTyrHisPheArgGluGlyAspLeuIleAlaValProThrGlyPheAla	146
QY	421	TTCTGCTCTACACACACGACACTGATGTGTGCTGTTCTTCTTACTGACACCAAC	480
DB	147	TyrTrpMetTyrAsnAsnGluAspThrProValValAlaValSerLeuIleAspThrAsn	166
QY	481	NACAAACGACACGAGCTTCAGTTCCCGCAGGAGATTCATTTGGCTGGGAACGAG	540
DB	167	SerPheGlnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu	186
QY	541	CAAGAGTTCTTAAGGTACCAAGCAACAAAGCAGACCAAGCAGACGAGAGCTTACCATAT	600
DB	187	GlnGluPheLeuGlnTyrGlnProGlnLysGlnGlnGly-----	199
QY	601	AGCCCATACAGCCCGCAAAAGTCAGCCTAGACAAAGAGCGGTGAATTTAGCCCTCGAGA	660
DB	200	-----Gly	200

QY	661	CAGCACAGCCGACAGAAACGAGCAGGACAAAGAAGAAGAAAAACGAAGGTGGAACATCTTC	720
DB	201	ThrGlnSerGlnLysGlyLysArgGlnGlnGluGluGluAsnGluGlySerIleLeu	220
QY	721	AGCGCTTCACGCCGAGTTCTCTGAAACAAGCTTCCAGGTTGACGACAGACAGATAGT	780
DB	221	SerGlyPheAlaProGluPheLeuGluHisAlaPheValVal---AspArgGlnIleVal	239
QY	781	CAAAACCTTAAGAGCGAGCAGCAGAGTAGTGAAGAAGAGGAGCCATTGTGACAGTGAAGGA	840
DB	240	ArgLysLeuGlnGlyGluAsnGluGluGluGlyAlaIleValThrValLysGly	259
QY	841	GGCTCAGAACTCTTGAGCCCA-----GATAGAAAGACGCTGCCGACCAAGAAGAG	891
DB	260	GlyLeuSerValIleSerProThrGluGlnGlnGlnArgProGluGluGlu	279
QY	892	GAATACGATGAAGATGAATATGATACGATGAAGAGGATAGAGCGCTGGCGGGGAGC	951
DB	280	LysProAsp-----CysAspGluLysAspLysHisCysGlnSerGlnSer	294
QY	952	AGAGCAGGGGGAATGTATTGAAGAGACGATCTGCACCGCAAGTGTAAAGAAACATC	1011
DB	295	Arg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgHisAsnIle	311
QY	1012	GGTAGAAACAGATCCCTGACATCTACAACCTCAAGCTGCTCACTCAAACTGCCAAC	1071
DB	312	GlyGlnThrSerSerProAspIlePheAsnProGlnAlaGlySerIleThrThrAlaThr	331
QY	1072	GATCTCAACCTTCTAATACTTAGTGGCTTGACCTAGTCTGATGATATGGAATCTCTAC	1131
DB	332	SerLeuAspPheProAlaLeuSerTrpLeuLysLeuSerAlaGlnPheGlySerLeuArg	351
QY	1132	AGGAATGATTGTTTGTCTCCTACCTACCAACACCAACGACACAGCATCATATATCGATTG	1191
DB	352	LysAsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleIleTyrAlaLeu	371
QY	1192	AGGGACCGGCTCAGCTGCAAGTCTGACACAGCAACCGCAACAGAGTGTACGACGAGGAG	1251
DB	372	AsnGlyArgAlaLeuValGlnValAlaAsnCysAsnGlyGluArgValPheAspGlyGlu	391
QY	1252	CTTCAAGAGGTGACGCTGCTTGTGTGTCACAGAACTTCGCGCTCGCTGGAAGTCCGAG	1311
DB	392	LeuGlnGluGlyGlnValLeuIleValProGlnAsnPheAlaValAlaAlaArgSerGln	411
QY	1312	AGCGAAGACTTCGAATAGTGTGCTCAAGACAGACTCAAGSCCCAGCATAGCCCACTC	1371
DB	412	SerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsnLeu	431
QY	1372	GCCGTGAAAACTCCGTCATAGATAACCTCGCGGAGGAGGTGGTTGCAAAATTCATATGC	1431
DB	432	AlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnGlnThrPheAsn	451
QY	1432	CTCCAAAGGAGCAGGCAAGCAGCTTAAGAACAAACACCCCTTCAGTTCTTCGTT---	1488
DB	452	LeuArgGlnSerGlnValSerGluLeuLysTyrGluGlyAsnTrpGlyProLeuValAsn	471
QY	1489	CCACCGTCTCAGCAG---TCTCCGAGG	1512
DB	472	ProGluSerGlnGlnGlySerProArg	480

RESULT 11
 S14393
 Legumin A2 precursor - fava bean
 C:Species: Vicia faba (fava bean)
 C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
 C:Accession: S14393
 R:Schlesier, B.; Bassuener, R.; van Hal, N.; Muentz, K.
 Nucleic Acids Res. 18, 7146, 1990
 A:Title: The cDNA derived primary structure of two distinct legumin A subunit precursors
 A:Reference number: S14392; MUID:91088307; PMID:2263481
 A:Accession: S14393
 A>Status: preliminary
 A:Molecule type: mRNA

A;Residues: 1-500 <SCH>
A;Cross-references: UNIPROT:Q99304; UNIPARC:UPI00000A2504; EMBL:X55014; NID:g22007; PIDN
C;Superfamily: glycinin

Alignment Scores: 4.22e-97 Length: 500
Pred. No.: 1414.50 Matches: 280
Score: 67.3% Conservative: 76
Percent Similarity: 52.9% Mismatches: 100
Best Local Similarity: 50.9% Indels: 73
Query Match: 2 Gaps: 8
DB:

US-10-728-323-3 (1-1524) x S14393 (1-500)

```

QY 1 CGGCAGACCGGAGAGAACGGTGCCTCAATCCAGCGCTCAATCGCGAGACCTGAC 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 23 ArgGluGlnSerGlnGlnAsnGluCysGlnLeuGluArgLeuAspAlaLeuGluProAsp 42
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 AATCGCATTCATCAGAGCGGCTTACATTGAGACTTGGAAACCCCAACACAGGAGTTC 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 43 AsnArgileGluSerGluGlyLeuileGluThrTrpAsnProAsnAsnArgGlnPhe 62
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCCGCGCAACGCCCTTCGTAGCCCT 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 63 ArgCysAlaSerValAlaLeuSerArgAlaThrLeuGlnArgAsnAlaLeuArgPro 82
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 181 TTCTACTCCAACTCCCGAGGAGATCTTCATCCAGCAAGGAAGGGATCTTTGGTTG 240
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 83 TyrTy-SerAsnAlaProGlnGluileTyrIleGlnGlnGlyAsnGlyTyrPheGlyMet 102
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 241 ATATTCCCTGGTTCTCTAGACACTATGAAGAGCCTTCACACAAAGTCTCGATCTCAG 300
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 103 ValPheProSerCysProGluThrPheGluGluPro-----GlnGln 116
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 301 TCCCAAGACCAACCAAGACGCTCTCCAGGAGAGAGACCAAGCCCAACAGCAAGATAGT 360
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 117 SerGlu-----GlnGlyGlu-----GlyGlyArgTyrArgAspSer 128
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 361 CACCAGAGGTGCACCGTTTCGATGAGGGTGTCTCATTTGCACTTCCCAACCGGTGTGCT 420
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 129 HisGlnLysValAsnArgPheArgGluGlyAspIleAlaValProThrGlyIleVal 148
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 421 TTCTGCTCTACACACACGACACTGTATGTTGCTGCTTCTCTTACTGACCAAC 480
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 149 PheTrpMetTyrAsnAspGlnAspThrProValIleAlaIleSerLeuThrAspIleGly 168
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 481 AACACGACACACAGCTTGTATCAGTTCCCGAGAGATTCAATTTGGCTGGGAACACGGAG 540
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 169 SerSerAsnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 188
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 541 CAAGAGTTCTTAAGGTACCGACAAACAAAGCAGACAAAGCAGAAAGCTTACCATAT 600
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 189 GlnGluPheLeuArgTyrGlnHisGln-----GlnGlnGlnGlnGlnGlnGln 197
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 601 AGCCCATACAGCCCGCAAGTACAGCTAGCTAGACAGAGAGCGTGAATTAGCCCTCGAGGA 660
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 197 -----GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 197
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 661 CAGCACAGCCGACAGAACGAGCAGGACAGAAAGAAAGAAAGGTGGAAACATCTTC 720
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 198 -----GlnGlyGlyLysGluGluGlnAspAsnAspGlyAsnAsnIlePhe 212
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 721 AGCGGTTTACGCGGAGTTCTTGGAACAAGCTTCCAGGTTGACGACAGAGATGTG 780
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 213 SerGlyPheLysArgAspPheLeuGluAspAlaLeuAsnVal---AsnArgHisIleVal 231
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 781 CAAACCTAGACGAGACCGAGAGTGAAGAGAGGAGCCATTGTGACAGTGAAGGGA 840
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 232 AspArgLeuGlnGlyArgAsnGluAspGluGluLysGlyAlaIleValLysValLysGly 251
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 841 GGCTCTCAGATCTTTCAGCCCGACATAGAAAG-----AGACGTGCGCGAGAA 885
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 252 GlyLeuSerIleIleThrProProGluArgGlnAlaArgHisProArgGlySerArgGln 271
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

886 GAAGAGGAATACGATCAAGATCAATATGAA----- 915

```

Db 272 GluGluAspGluAspGluAspGluArgGlnProSerHisLysSerArg 291
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 916 -----TACGATGAAGAGGATAGA-----AGCGTGGCAGGGGAACGACA 954
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 292 ArgGlyGluAspGluAspAspLysGluLysArgHisSerGlnLysGlyGluSerArgArg 311
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 955 GGCAGGGGAATGGTATTGAAGACGATCTCCACCGCAAGTGTAAAGAACATTTGTT 1014
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 312 HisGlyAspAsnGlyLeuGluThrValCysThrAlaLysLeuArgLeuAsnIleGly 331
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1015 AGAAACAGATCCCTCGACATCTACACCTCAAGCTGCTCACTCAAACTCCCAACGAT 1074
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 332 SerSerSerSerProAspIleTyrAsnProGlnAlaGlyArgIleLysThrValThrSer 351
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1075 CTCACCTTCTAATCTAGTGGCTTGGACCTAGTCTGTAATATGGAATCTCTACAGG 1134
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 352 LeuAspLeuProValLeuArgTrpLeuLysLeuSerAlaGluHisGlySerLeuArgLys 371
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1135 AATGATCTTTTGTCTCATCAACACCAACGACACAGCATCATATATCGATTGAGG 1194
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 372 AsnAlaMetPheValProHisTyrAsnLeuAsnAlaAsnSerIleLeuTyrAlaLeuLys 391
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1195 GACGGGCTCAGTCCAGTCTGGACACGACGCAAGCTGTACGACGAGGAGCTT 1254
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 392 GlyArgAlaArgLeuGlnValValAsnCysAsnGlyAsnThrValPheAspGluGluLeu 411
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1255 CAAGAGGGTCACTGCTGTGTCGACACAGAACTTCGCGCTGCTGAAAGTCCCAAGAGC 1314
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 412 GluAlaGlyArgAlaLeuThrValProGlnAsnTyrAlaValAlaLysSerLeuSer 431
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1315 GAAACTTCGAATACGTGGCATTCAGACAGACTCAAGCCCGACGATACGCCCTCGCC 1374
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 432 AspArgPheThrTyrValAlaPheLysThrAsnAspArgAlaGlyIleAlaArgLeuAla 451
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1375 GGTGAAACTCCGTCATAGATAACCTGCGGAGGAGGTGGTTCGCAATTCATATGCCCTC 1434
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 452 GlyThrSerSerValIleAsnAspMetProValAspValValAlaAlaThrPheAsnLeu 471
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1435 CAAAGGAGCAGCAGGAGGAGCTTAAGAAACAACACCCCTTCAAGTTCCTTCACCCG 1494
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 472 GluArgAsnGluAlaArgGlnLeuLysSerAsnAsnProPheLysPheLeuValProPro 491
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1495 TCTCAGCAGCTCTCCGAGGCGTGTGGCT 1521
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 492 ArgGluSerGlnLysArgAlaSerAla 500
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 12

T06452

probable legumin A precursor - garden pea

C;Species: Pisum sativum (garden pea)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-Oct-2004

C;Accession: T06452

R;Lycett, G.W.; Croy, R.R.D.; Shirsat, A.H.; Richards, D.M.; Boulter, D.

Nucleic Acids Res. 13, 6733-6743, 1985

A;Title: The 5'-flanking regions of three legumin genes: comparison of the DNA sequence

A;Reference number: Z15687; MUID:86041868; PMID:2997721

A;Accession: T06452

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-507 <byC>

A;Cross-references: UNIPARC:UPI000016DF29; EMBL:X02982; NID:g20777; PIDN:CAA26720.1; PI:1

C;Genetics:

A;Intons: 96/1; 179/3; 388/3

C;Keywords: seed; storage protein

F;1-21/Domain: signal sequence #status predicted <SIG>

F;22-332/Product: legumin A, alpha chain #status predicted <ALP>

Alignment Scores:

Pred. No.: 7.79e-95 Length: 507

Score: 1384.00 Matches: 283

Percent Similarity:	63.9%	Conservative:	66
Best Local Similarity:	51.8%	Mismatches:	97
Query Match:	49.8%	Indels:	100
DB:	2	Gaps:	9

US-10-728-323-3 (1-1524) x T06452 (1-507)

QY	1	CGGCAGCAACCGGAGGAGAACCGCGTGCACGATTCCTCAGCCCTCAATGCGCAGACACCTGAC	60
Db	23	ArgGluInProGlnGlnAsnGluCysGlnLeuGluArgLeuAspAlaLeuGluProAsp	42
QY	61	AATCGCATTTCAATTCAGAGCGCGGTTCACATTCAGACTTCGAAACCCCAACACAGGAGTTC	120
Db	43	AsnArgIleGluSerGluGlyGlyLeuIleGluThrTrpAsnProAsnAsnLysGlnPhe	62
QY	121	GAATGGCCCGCGTGCCTCTCTCGCTTAGTCTCCGCGCAACCGCCCTTCGTAGGCCT	180
Db	63	ArgCysAlaGlyValAlaLeuSerArgAlaThrLeuGlnArgAsnAlaLeuArgArgPro	82
QY	181	TTCTACTCCAAATGCTCCCGAGAGATCTTCATCAGCAAGGAGGAGGATATCTTTGGGTG	240
Db	83	TyrTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyTyrPheGlyMet	102
QY	241	ATATTCCTCGTGTCTTAGACACTATGAAGAGGCTCTACACAAAGTCTGTCATCTCAG	300
Db	103	ValPheProGlyCysProGluThrPheGluGluPro	116
QY	301	TCCCAAGACCAAGAGCGTCTCCAAGAGAGAACCAAGCCCAACGACCAACGAGATAGT	360
Db	117	SerGlu	128
QY	361	CACCAAGAGGTCCACCGTTTCGATGAGGTGATCTCATTCAGTCTCCACCGGTGTGCT	420
Db	129	HisGlnLysValAsnArgPheArgGluGlyAspIleAlaValProThrGlyIleVal	148
QY	421	TTCTGGCTCTACAACGACCAACGACTGATGTTGCTGTTCTCTTACTGACACCAAC	480
Db	149	PheTrpMetTyrAsnAspGlnAspThrProValIleAlaValSerLeuThrAspIleArg	168
QY	481	AACCAACGACCAACGAGTTCAGTTCCTCCAGGAGATTCAAATTTGGCTGGGAACGAG	540
Db	169	SerSerAsnAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnHisGlu	188
QY	541	CAAGAGTCTTAAAGTTACAGCAACAAGACGACAAAGACGACGAAAGCTTACCATAF	600
Db	189	GlnGluPheLeuGlnTyrGlnHisGln	197
QY	601	AGCCCATACAGCCCGCAAAAGTCAGCCTAGACAAGAAGCGTGAATTTAGCCCTCGAGGA	660
Db	197		197
QY	661	CAGCAAGCCGAGAGAACGACGACGAGCAAGAAGAAGAAACGAAGTGGAAACATCTTC	720
Db	198		212
QY	721	AGCGCTTACCGCGGATTCCTGGAACAGCCCTCCAGGTTGACGACGACAGATAGTG	780
Db	213	SerGlyPheLysArgAspTyrLeuGluAspAlaPheAsnVal	231
QY	781	CAAAACCTAAGAGCGGACCGCAGTAGTGAAGAAGAGGAGCCATTTGTGACAGTACGAGGGA	840
Db	232	AspArgLeuGlnGlyArgAsnGluAspGluGluLysGlyAlaIleValLysValLysGly	251
QY	841	GGCTCAGAAATCTTACGCGCCAGATAGAAG	870
Db	252	GlyLeuSerIleIleSerProProGluLysGlnAlaArgHisGlnArgGlySerArgGln	271
QY	871		891
Db	272	GluGluAspGluAspGluGluLysGlnProArgHisGlnArgGlySerArgGlnGluGlu	291
QY	892	GAATACGATGAAGATGAATATATGAATACGATGAAGAGGATAGAAGCGGTGCG	942

Db	292	GlulaspGluaspGluGluAurGlnPröärHisGlnAargArgGlyGluGluGlu 311
Qy	943	-----AGGGGAAGCAGAGCGAGGGG----- 963
Db	312	GlulaspLysLysGluAurGlyGlySerGlnLysGlyLysSerArgAsgGlnGlyAsp 331
Qy	964	AATGGTATTGAAGAGCATCTCGACCGCAAGTGCTAAAAAGAACAATTGGTAGAAACAGA 1023
Db	332	AsnGlyLeuGluGluThrValCysThrAlaLysLeuArgLeuAsnIleGlyProSerSer 351
Qy	1024	TCCCTCAGCATCTACAACCTCTCAAGCTGGTTCTACTCAAACTGCCAACGATCTCAACCTT 1083
Db	352	SerProAspIleTyrAsnProGluuAlaGlyArgIleLysThrValThrSerLeuAspLeu 371
Qy	1084	CTAATACTTAGTGGCTTGGACCTAGTGCCTGAATATATGAAATCTCTACAGGAATGCAATTG 1143
Db	372	ProValLeuArgTrpLeuLysLeuSerAlaGluHisGlySerLeuHisLys----- 388
Qy	1144	TTTGTCTCTACTACAACACCAACGACACAGCATCATATATCGATTGAGGAGCGGCT 1203
Db	389	-----AsnAlaAsnSerIleIleTyrAlaLeuLysGlyArgAla 401
Qy	1204	CAGTGTCAAGTCGTGGACAGCAACGGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGGT 1263
Db	402	ArgLeuGlnValValAsnCysAsnGlyAsnThrValPheAspGlyGluLeuGluAlaGly 421
Qy	1264	CACGTGCTTGTGTGCACAGAACTTCGCGCTCGCTGGAAGTCCCAAGCGAGAACTTC 1323
Db	422	ArgAlaLeuThrValProGlnAsnTyrAlaValAlaLysSerLeuSerAspArgPhe 441
Qy	1324	GAATACGTGGCATTCAGACAGACTCAAGCCCGACAGTACGCCAACCCTCGCGCGGTGAAAC 1383
Db	442	SerTyrValAlaPheLysThrAsnAspArgAlaGlyIleAlaArgLeuAlaGlyThrSer 461
Qy	1384	TCCGTCTATAGTAACCTCGCGAGGAGTGGTTCGCAAAATTCATATGCCCTCCAAAGGGAG 1443
Db	462	SerValIleAsnAsnLeuProLeuAspValValAlaAlaThrPheAsnLeuGlnArgAsn 481
Qy	1444	CAGCGAAGGAGCTTAAAGAACAAACCCCTTCAAAGTTCCTCGTTCACCGTCTCACGAC 1503
Db	482	GluAlaArgGlnLeuLysSerAsnAsnProPheLysPheLeuValProAlaAaArgGluSer 501
Qy	1504	TCTCCGAGGCTGTGGCT 1521
Db	502	GluAsnArgAlaSerAla 507

RESULT 13

Sl4392

legumin A1 precursor - fava bean (fragment)

C:Species: Vicia faba (fava bean)

C>Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 05-Oct-2004

C:Accession: S14392

R:Schlesier, B.; Bassuener, R.; van Hai, N.; Muentz, K.

Nucleic Acids Res. 19, 7146, 1990

A:Title: The cDNA derived primary structure of two distinct legumin A subunit p

A:Reference number: S14392; MUID:91088307; PMID:2263481

A:Accession: S14392

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-497 <SCH>

A:Cross-references: UNIPROT:Q03971; UNIPARC:UPI000000AOEEA; EMBL:X55013; NID:922

Alignment Scores:

Pred. No.:	3,62e-94	Length:	497
Score:	1375.00	Matches:	276
Percent Similarity:	68.8%	Conservative:	79
Best Local Similarity:	53.5%	Mismatches:	107
Query Match:	49.5%	Indels:	54
DB:	2	Gaps:	11

US-10-728-323-3 (1-1524) x S14392 (1-497)

Qy 25 TGGCAGTTCACGGCGCTCAATGGCGCAGACCTGCAATCGCATTTGAATCAGAGGGCGGT 84

```
Db      1  CysGlnLeuGluArgLeuAspAlaLeuGluProAspAsnArgIleGluSerGluGlyGly 20
      85  TACATTGAGACTTTGGAACCCCAACACACGAGAGTTCGAATTCGCGCGCGCTCGCCCTCTCT 144
Db      21  LeuIleGluThrTrpAsnProAsnAsnArgGlnPheArgCysAlaArgValAlaLeuSer 40
      145  CGTTTAGTCTCCGCGCGCAACGCGCTTCGTAGCCCTTCTACCTCAATGCTCCCGAGGAG 204
Db      41  ArgAlaThrLeuGlnArgAsnAlaLeuArgArgProTyrTyrSerAsnAlaProGlnGlu 60
      205  ATCTTCATCAGCAAGGAAGGATACTTTGGTTCATATCCCTGGTGTCTCTACACAC 264
Db      61  IleTyrIleGlnGlnGlyAsnGlyTyrPheGlyMetValPheProSerCysProGluThr 80
      265  TATGAAGAGCTCACACAAAGTCTGTCGATCTCAGTCCCAAGACACCAAGACGCTCTC 324
Db      81  PheGluGluPro-----GlnGlnSerGlu----- 88
      325  CAAGGAGAGACCAAGCCAAACGACGAGTAGTCACAGAAAGGTGCACCGTTTCGAT 384
Db      89  GlnGlyGlu-----GlyArgArgTyrArgAspSerHisGlnLysValAsnArgPheArg 406
      385  GAGGGTATCTCATTCGAGTCCACCGGTGTTGCTTCTCGCTCTACACGACACACGAC 444
Db      107  GlnGlyAspIleIleAlaValProThrGlyIleValPheTrpMetTyrAsnAspGlnAsp 426
      445  ACTGATGTGTTGCTGCTTCTTCTTACTGACCAACAAACACACGACACCAAGCTTGATCAG 504
Db      127  IleProValIleAlaIleSerLeuThrAspThrGlySerSerAsnAsnGlnLeuAspGln 146
      505  TTCCCGCAGGAGATTCAAATTGGCTGGGAACACGAGCAGAGAGTCTTTAAGGTACCAACG 564
Db      147  MetProArgArgPheTyrLeuAlaGlyAsnGlnGluGlnGluPheLeuArgTyrGlnHis 166
      565  CAAGCAGACACAAAGCAGACGAAAGCTTACCATATAGCCCATACACCGCGCAAGTCCAG 624
Db      167  GlnGlnGlyValLysGluGlu-----GlnAspAsn 176
      625  CCTAGACAAGAGCGTGAATTTAGCCCTCGAGGACAGCACACGCGCCGACGAGACGACGA 684
Db      177  AspGlyAsnGlnGlnGluPheLeu---ArgTyrGlnHis-----ArgGlnGlyVal 193
      685  GGCACAGAGAAGAAACGAAAGTGGAAACATCTTCAGCGGCTTCAGCGCGGAGTTCCTG 744
Db      194  LysGluGluGlnAspAsnAspGlyAsnAsnIlePheSerGlyPheAsnArgAspPheLeu 213
      745  GAACAAGCCTTCACAGTTGACGACACACATAGTGCAAAACCTTAAGAGCGGAGCCGAG 804
Db      214  GluAspAlaPheAsnVal---AsnArgHisIleValAspArgLeuGlnGly-----Arg 230
      805  AGTGAAGAAGAGGAGCCATTGTGACAGTGGAGGGAGCGCTCAGAAATCTTTGAGCCCGAT 864
Db      231  AsnGluGluArgGlyAlaIleValLysValLysGlyGlyLeuSerIleIleThrProPro 250
      865  AGAAAGAGA-----CGTGCCGACGCAAGAGGAATACGATGAA 903
Db      251  GluArgGlnAlaArgHisProArgLysArgGlnGluGluAspGluAspGlu 270
      904  GATGATATATGAA-----TACGATGAGAG 927
Db      271  AspGluLysGluGluArgGlnProSerHisIleLysSerArgArgGlyGluAspGluAsp 290
      928  GATAGA-----AGCGGTGGCAGGGGAAGCAGACGCGCGGGAATGTATT 972
Db      291  AspLysGluLysArgHisSerGlnLysGlyGluSerArgArgHisGlyAspAsnGlyLeu 310
      973  GAAGACGATCTCCACCGCAAGTCTAAAGAACATTCGTAGAAACAGATCCCTGAC 1032
Db      311  GluGluThrValCysThrAlaLysLeuArgLeuAsnIleGlySerSerSerProAsp 330
      1033  ATCTCAACCTCAAGCTGGTTCATCAAACTGCCCAACGATCTCAACCTTCTTAATCTT 1092
```

```
Db      331  IleTyrAsnProGlnAlaGlyArgIleLysThrValThrSerValAspLeuProValLeu 350
      1093  AGGTGGCTTGGACCTAGTGTCTGAATATGGAATCTCTACAGGAATGCATTTGTTCGCT 1152
Db      351  ArgTrpLeuLysLeuSerAlaGluHisGlySerLeuArgLysAsnAlaMetPheValPro 370
      1153  CACTACACACCAACGACACACATCATATATCGATTGAGGGGCGGCTCAGGTGCAA 1212
Db      371  HisTyrAsnLeuAsnAlaAsnSerValLeuTyrAlaLeuLysGlyArgAlaArgLeuGln 390
      1213  GTCTGGACAGCAACGGCAACAGAGTACGACGAGGAGCTTCAAGAGGGTCACGTGCTT 1272
Db      391  ValValAsnCysAsnGlyAsnThrValPheAspGlyGluLeuGluAlaGlyArgAlaLeu 410
      1273  GTGTGCCACAGAACTTCCCGCTCGTGGAAAGTCCGACGAGGAGGAGTCCGAATACGTG 1332
Db      411  ThrValProGlnAsnTyrValValAlaAlaLysSerLeuSerAspArgPheThrTyrVal 430
      1333  GCATTCACAGACACTCAAGGCCAGCATAGCAACCTCGCGGTGAAAATCCCGTCATA 1392
Db      431  AlaPheLysThrAsnAspArgAlaGlyIleAlaIaIa-gLeuAlaGlyThrSerSerValIle 450
      1393  GATAACTCCCGAGAGGTGTTGCAAAATTCATATGGCTCCAAAGGAGGAGGAGCAAG 1452
Db      451  AsnAspLeuProLeuAspValValAlaAlaThrPheAsnLeuGluArgAsnGluAlaArg 470
      1453  CAGCTTAAGAAACAACCCCTTCAGTTCCTTCGTTCCACCGTCTCAG 1500
Db      471  GlnLeuLysPheAsnAsnProSerArgPheLeuValProProArgGlu 486

RESULT 14
S51941
prunin 1 precursor - almond
N;Alternate names: salt-soluble globulin; seed storage protein
C;Species: Prunus dulcis (almond)
C;Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 05-Oct-2004
R;Garcia-Mas, J.; Messeguer, R.; Arus, P.; Puigdomenech, P.
Plant Mol. Biol. 27, 205-210, 1995
A;Title: Molecular characterization of cDNAs corresponding to genes expressed during al
A;Reference number: S51940; MUID:95170003; PMID:7865791
A;Accession: S51941
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-551 <AR>
A;Cross-references: UNIPROT:Q43607; UNIPARC:UPI0000177DF8; EMBL:X78119
A;Note: the source is designated as Prunus amygdalus
R;Garcia-Mas, J.; Messeguer, R.; Arus, P.; Puigdomenech, P.
submitted to the EMBL Data Library, March 1994
A;Description: Molecular characterization of cDNAs corresponding to proteins related to
A;Reference number: S42473
A;Accession: S42474
A;Molecule type: mRNA
A;Residues: 1-60, 'G', 62-551 <GAW>
A;Cross-references: UNIPARC:UPI00000ACDD6; EMBL:X78119; NID:g460805; PIDN:CAA55009.1; P
C;Keywords: seed; storage protein
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-367/Product: prunin 1 alpha chain #status predicted <MAT1>
F;368-551/Product: prunin 1 beta chain #status predicted <MAT2>
F;108-374/Disulfide bonds: #status predicted

Alignment Scores:
Pred. No.: 9,02e-72 Length: 551
Score: 1073.50 Matches: 228
Percent Similarity: 57.2% Conservative: 89
Best Local Similarity: 41.2% Mismatches: 160
Query Match: 38.6% Indels: 77
DB: 2 Gaps: 11

US-10-728-323-3 (1-1524) x S51941 (1-551)
QY 4 CAGCAACCGGAGGAGAACGCGTGCCAGGTTCCAGCGCTCAATGCGCAGACGCTGACAAT 63
```



```
QY 79 GCGGTTCATTGAGACTGGACCCCAACACACAGGAGTTCGAATGCGCGCGCTGCC 138
Db 79 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
50 GlyGlyLeuIleGlnThrPheAsnSerGlnHisProGluLeuLysCysAlaGlyValThr 69
QY 139 CTCCTCGCTAGTCTCCGCGCCCAACAGCCCTTCGTAGGCTTCTTACTCCCAATGCTCC 198
Db 139 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
70 ValSerLysLeuThrLeuAsnArgAsnGlyLeuHisLeuProSerTyrSerProTyrPro 89
QY 199 CAGGAGATCTTCATCAGCAAGGAAGGGATACCTTTGGTGTGATATCCCTGTTGCTCT 258
Db 199 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
90 ArgMetIleIleAlaGlnGlyLysGlyAlaLeuGlyValAlaIleProGlyCysPro 109
QY 259 AGACACTATCAAGAGCTCACACACAGGTGCTGATCTCAGTCCCAAGACACCAAGA 318
Db 259 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
110 GluThrPheGluGluProGlnGluGlnSerAsnArgGlySerArg----- 125
QY 319 CGTCTCCAAGGAGACCAAGCCCAACAGACAGATAGTACCAGAAAGGTGCACCGT 378
Db 319 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
126 -----SerGlnLysGlnGlnLeuGlnAspSerHisGlnLysIleArgHis 140
QY 379 TTCATGAGGGTGAATCTCATGTGAGTTCACCGGTGTTGCTTCTTACACAGAC 438
Db 379 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
141 PheAsnGluGlyAspValLeuValIleProProGlyValProTyrTrpThrTyrAsnThr 160
QY 439 CAGCAGACTGATGTTCTGCTGTTCTCTTACTGACACCAACACACACAGCAGCTT 498
Db 439 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
161 GlyAspGluProValValAlaIleSerLeuLeuAspThrSerAsnPheAsnAsnGlnLeu 180
QY 499 GATCAGTTCCTCCAGGAGATTCATTTGGTGGAAACACGAGCAAGAGTTCCTTAAGGTAC 558
Db 499 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 AspGlnThrProArgValPheTyrLeuAlaGlyAsnProAspIleGluTyr----- 197
QY 559 CAGCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATATAGCCCATACGCCGCAA 618
Db 559 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
198 -----ProGluThrMetGln 202
QY 619 AGTCAGCCTAGACAAGNAGCGTGAATTTAGCCCTCGAGGACAGACACGCCCGCAGAA 678
Db 619 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
203 GlnGlnGlnGlnGlnLysSerHisGlyArgLysGlnGlnHisGln----- 219
QY 679 CGAGCAGGACAAGAAGAAACGAAAGGTGGAACATCTTCAGCGGCTTCACGCCGAG 738
Db 679 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
220 -----GlnGluGluGluGluGlySerValLeuSerGlyPheSerLysHis 236
QY 739 TTCCTGAAACACGCTTCAGGTTCAGCAGACAGATAGTGCACAACTTAAGCGGAG 798
Db 739 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
237 PheLeuAlaGlnSerPheAsnThrAsnGlu--AspIleAlaGluLysLeu-----Gln 253
QY 799 ACCGAGAGTCAAGAAGGGAGCCATTGTGACAGTGAAGGGAGGCTCAGAAATCTTGAGC 858
Db 799 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
254 SerProAspAspGluArgLysGlnIleValThrValGluGlyLeuSerValIleSer 273
QY 859 CCATAGAGAAGAGACTGCGCAGAAAGAGGAAATACCATGAA----- 903
Db 859 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
274 ProLysTrpGlnGlnGlnAspGluAspGluAspGluAspGluAspGluAspGlu 293
QY 904 -----GATGAATAT 912
Db 904 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
294 GlnIleProSerHisProProArgProSerHisGlyLysArgGluGlnAspGluAsp 313
QY 913 GAATACGATGAAGGATAGA--AGCGTGGCAGGAGGAGCAGAGC----- 957
Db 913 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
314 GluAspGluAspGluAspLysProArgProSerArgProSerGlnGlyLysArgGluGln 333
QY 957 ----- 957
Db 334 AspGlnAspGlnAspGluAspGluAspGluAspGlnProArgLysSerArgGlu 353
QY 958 -----AGGGG 963
Db 354 TrpArgSerLysLysThrGlnProArgProArgProArgGlnGluProArgGluArgGly 373
QY 964 -----AATGTTATTGAAGACGATCTGCACCGCAAGTGTCTAAAGAACATT 1011
```

```
Db 374 CysGluThrArgAsnGlyValGluGluAsnIleCysThrLeuLysLeuHisGluAsnIle 393
QY 1012 GTTGAACACAGATCCCTCGACATCTACAACCTCAAGCTGAGTGGTCTACTCAAACTGCCAAC 1071
Db 394 AlaArgProSerArgAlaAspPheTyrAsnProLysAlaGlyArgIleSerThrLeuAsn 413
QY 1072 GATCTCAACCTTCTAATACTAGTGGCTTGGACCTAGTGGCTGAATATGAAATCTCTAC 1131
Db 414 SerLeuThrLeuProAlaLeuArgGlnPheGlnLeuSerAlaGlnIleValValLeuTyr 433
QY 1132 AGCAATGCAATGCTTGTCTCCTCACTACACACCAACGACACAGCATCATATATCGATTG 1191
Db 434 LysAsnGlyIleTyrSerProHisTrpAsnLeuAsnAlaAsnSerValIleTyrValThr 453
QY 1192 AGGGACGGGCTCAGGTGCAAGTCTGTGGACACAGCAACGCGCAACAGAGTGTACACGAGGAG 1251
Db 454 ArgGlyGlnGlyLysValArgValValAsnCysGlnGlyAsnAlaValPheAspGlyGlu 473
QY 1252 CTTCAAGAGGCTCAGTGTCTGCTGCTGACAGAACTTCGCGTGGTGGAAAGTCCCGAG 1311
Db 474 LeuArgArgGlyGlnLeuLeuValProGlnAsnPheValValAlaGluGlnAlaGly 493
QY 1312 AGCGAAGACTTCGAATACGTGCAATTCAGACAGACTCAAGGCCCAAGTAGCAACCTC 1371
Db 494 GluGlnGlyPheGluTyrIleValPheLysThrHisHisAsnAlaValThrSerTyrLeu 513
QY 1372 GCGGTGAAACTCCGCTCATAGATACTGCGGAGGAGGTGGTTCGCAAAATTCATATGGC 1431
Db 514 -----LysAspValPheArgAlaIleProSerGluValLeuAlaHisSerTyrAsn 530
QY 1432 CTCCAAGGAGCAGCAGGAGGAGCTTAAAGAACCAACACCCCTTCAAGTTCCTCGTTCCA 1491
Db 531 LeuArgArgGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuValPro 550
QY 1492 CCGTCTCAGCAGCTCTCCGAGGCTGTGGCT 1521
Db 551 ProLysGluSerGlnArgValValAla 560
```

Search completed: May 15, 2006, 22:22:27
Job time : 80.0021 secs

GenCore version 5.1.1.8
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: May 15, 2006, 21:23:21 ; Search time 69.4413 Seconds
(without alignments)
4645.181 Million cell updates/sec

Title: US-10-728-323-3

Perfect score: 2779

Sequence: 1 cggcgacacaggagagaa.....ctccgagggtgtggtcttaa 1524

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q-/abss/ABSSWEB_spool/US10728323/runat_15052006.172132.22377/app_query.fasta_1
-DB=uniprot -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abss02p
-USER=US10728323 @CGN 1.1.694 @runat 15052006.172132.22377 -NCPU=6 -ICPU=3
-NO MAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Uniprot 05.80:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2673	96.2	507	082580 ARAHY	082580 arachis hyp
2	2579.5	92.8	531	Q516T2 ARAHY	Q516t2 arachis hyp
3	2479	89.2	529	Q9FZ11 ARAHY	Q9fz11 arachis hyp
4	2455	88.3	536	Q647H4 ARAHY	Q647h4 arachis hyp
5	2451.5	88.2	537	Q647H3 ARAHY	Q647h3 arachis hyp
6	2444	87.9	538	Q8LKN1 ARAHY	Q8lkn1 arachis hyp
7	2436	87.7	536	Q672T4 ARAHY	Q672t4 arachis hyp
8	2421	87.1	530	Q9SQH7 ARAHY	Q9sqh7 arachis hyp
9	1834	66.0	510	Q61WG5 ARAHY	Q61wg5 arachis hyp
10	1651	59.4	484	Q647H2 ARAHY	Q647h2 arachis hyp
11	1502	54.0	481	GLYG3 SOYBN	GLYG3 soybean
12	1499.5	54.0	485	GLYG2 SOYBN	GLYG2 soybean
13	1499.5	54.0	485	Q549Z4 SOYBN	Q549z4 soybean
14	1496	53.8	495	GLG1 SOYBN	GLG1 soybean
15	1489	53.6	481	Q852U5 SOYBN	Q852u5 soybean
16	1488.5	53.6	482	Q852U4 SOYBN	Q852u4 soybean

17	1457.5	52.4	498	2	Q41702 VICSA	Q41702 vicia sativ
18	1455	52.4	520	1	LEGA2 PEA	P15838 pium sativ
19	1443	51.9	517	2	Q9T0P5 PEA	Q9t0p5 pium sativ
20	1440	51.8	517	1	LEGA PEA	P03857 pium sativ
21	1431.5	51.5	482	2	Q41676 VICNA	Q41676 vicia narbo
22	1414.5	50.9	500	2	Q99304 VICFA	Q99304 vicia faba
23	1375	49.5	497	2	Q03971 VICFA	Q03971 vicia faba
24	1186	42.7	496	2	Q9SMJ4 CICAR	Q9smj4 cicar ariet
25	1110.5	40.0	515	2	Q8W1C2 CORAV	Q8w1c2 corylus ave
26	1074.5	38.7	488	2	Q41128 QUERO	Q41128 quercus rob
27	1069.5	38.5	551	2	Q43607 PRUDU	Q43607 prunus dulc
28	1059	38.1	542	2	Q8LK20 CASCR	Q8lk20 cascanea cr
29	1049	37.7	560	2	Q9S9D0 SOYBN	Q9s9d0 glycine max
30	1034.5	37.2	457	2	Q8GZP6 ANAOC	P05692 pium sativ
31	1032	37.1	503	1	LEGJ PEA	P05692 pium sativ
32	999.5	36.0	566	2	O24294 PEA	O24294 pium sativ
33	998	35.9	485	2	Q41703 VICSA	Q41703 vicia sativ
34	996.5	35.9	484	1	LEG84 VICFA	P05190 vicia faba
35	986	35.5	517	2	Q39922 GLYSO	Q39922 glycine soj
36	986	35.5	517	2	Q7GC77 SOYBN	Q7gc77 glycine max
37	985	35.4	517	2	Q9SB12 SOYBN	Q9sb12 glycine max
38	981.5	35.3	564	2	Q43673 VICFA	Q43673 vicia faba
39	980	35.3	517	2	P93707 SOYBN	P93707 glycine max
40	980	35.3	517	2	P93708 SOYBN	P93708 glycine max
41	978.5	35.2	219	2	Q8LL03 ARAHY	Q8ll03 arachis hyp
42	975	35.1	563	2	Q9SB11 SOYBN	Q9sb11 glycine max
43	974	35.0	563	2	Q39921 GLYSO	Q39921 glycine soj
44	973.5	35.0	562	1	GLYG4 SOYBN	P03858 glycine max
45	968	34.8	504	2	Q43608 PRUDU	Q43608 prunus dulc

ALIGNMENTS

RESULT 1

082580 ARAHY
ID 082580 ARAHY PRELIMINARY; PRT; 507 AA.
AC 082580;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Glycinin (Fragment).
GN Name=Arach3;
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99146968; PubMed=10021462;
RA Rabinov P., Helm E.M., Stanley J.S., West C.M., Sampson H.A.,
RA Burks A.W., Bannon G.A.;
RT "Molecular cloning and epitope analysis of the peanut allergen Ara h
3.";
RL J. Clin. Invest. 103:535-542(1999).
DR EMBL; AF093541; AAC63045.1; -; mRNA.
DR HSSP; P04776; 1UD1.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin_1; 2.
DR PRINTS; PR00439; 11SGLOBULIN
DR PROSITE; PS00305; 11S_SEED_STORAGE; UNKNOWN_1.
FT NON TER 1
SQ SEQUENCE 507 AA; 58349 MW; F3FB38BC3CB82DED CRC64;

Alignment Scores:
Pred. No.: 3,298-178 Length: 507
Score: 2673.00 Matches: 507
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0

Score:	2579.50	Matches:	494
Percent Similarity:	97.6%	Conservative:	2
Best Local Similarity:	97.2%	Mismatches:	11
Query Match:	92.8%	Indels:	1
DB:	2	Gaps:	1

US-10-728-323-3 (1-1524) x Q516T2_ARAHY (1-531)

QY	1	CGCGAGCAACCGGAGAGAACGCTGCCAGTTCACAGCGCTCAATGCGCAGAGACTGAC	60
DB	24	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	43
QY	61	AATCGCATTTGAATCAGAGCGGCTTACATTGAGACTTGGAACCCCAACACAGAGTTC	120
DB	44	AsnArgIleGluSerGluGlyGlyTyrIleGluThrTriPAsnProAsnAsnGlnGluPhe	63
QY	121	GAATGCGCGCGCTCGCTCTCGCTTGTAGTTCCTCCGCGCAACGCCCTTCGTAGGCT	180
DB	64	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro	83
QY	181	TTCTACTCCAAATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGATACCTTTGGTTG	240
DB	84	PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu	103
QY	241	ATATTCCCTGTTGCTCTAGACACTATGAAGACCTCACACAGAGTTCGTCATCTCAG	300
DB	104	IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgSerGln	123
QY	301	TCCCAAGACCCAGAGAGCTCTCCAGGAGAGACCAAGCAACAGCAGATAGT	360
DB	124	SerGlnArgProProArgArgLeuGlnGlyGluAspGlnSerGlnGlnGlnAspSer	143
QY	361	CACCAAGAGTGCACCGTTTCATAGAGGTGATCTCATTCGAGTTCACACCGGTGTCT	420
DB	144	HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	163
QY	421	TTCTGGCTCTACAACGACCACGACACTGATGTGTGCTGTTCTTCTTACTGACCAAC	480
DB	164	PheTyrLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn	183
QY	481	ACACGCAACACAGCTTCAGTTCCTCCAGGAGATTCATTTGGCTGGGACGGAG	540
DB	184	AsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnHisGlu	203
QY	541	CAAGAGTCTTAAAGTACAGCAACAAAGCAGACAAAGCAGACGAAAGAGCTTACCAT	600
DB	204	GlnGluPheLeuArgTyrGlnGlnGlnSerArgSerGlnSerArgArgSerLeuProTyr	223
QY	601	AGCCCATACAGCCCGCAAAAGTCAGCTAGACAAAGAGCGGTGAATTTAGCCCTCAGGA	660
DB	224	SerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArgGly	243
QY	661	CAGCACAGCCGAGACGACGACGACGACGACGACGACGACGACGACGACGACGAC	720
DB	244	GlnHisSerArgArgGluArgAlaGlyGlnGlnGluGluAsnGluGlyGlyAsnIlePhe	263
QY	721	AGCGGTTCACGCGGAGTTCCTGGAACCAAGCTTCAGGTTCACGACACAGATAGTG	780
DB	264	SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal	283
QY	781	CAAAACCTAAGAGCGGAGACCGAGAGTGAAGAGGAGGAGCCATTGTGACAGTGAAGGA	840
DB	284	GlnAsnLeuArgGlyGluAsnGlnSerGluGluGluGlyAlaIleValThrValArgGly	303
QY	841	GGCCTCAGATCTTGAGCCAGATAGAACAGACGTCGCCGACGACGACGACGACGAC	900
DB	304	GlyLeuArgIleLeuSerProAspArgLysArgGlyAlaAspGluGluGluGluTyrAsp	323
QY	901	GAAGATGAATATGATACCATGAAGAGGATAGAAGCGGTGGCAGGGGAGAGGAGCAGG	960
DB	324	GluAspGluTyrGluTyrAspGluGluAspArgArgGlyArgGlySerArgGlyArg	343
QY	961	GGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAGAGAACATTCGTGAAC	1020

DB	344	GlyAsnGlyIleGluGluThrIleCysThrAlaSerValLysLysAsnIleGlyArgAsn	363
QY	1021	AGATCCCTGACATCTACAACCTCAAGCTTCACTCAAAACCTGCAACCATCTCAAC	1080
DB	364	ArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeuAsn	383
QY	1081	CTTCTAATACCTTAGTGCTTGAGACTAGTCTGAATATGGAATCTCTACAGGAATGCA	1140
DB	384	LeuLeuIleLeuAlaGlyTyrLeuGlyLeuSerAlaGluTyrGlyAsnLeuTyrArgAsnAla	403
QY	1141	TTGTTTTCGCTCACTACAACACCAACGACACACATATATATATATATATATATATAT	1200
DB	404	LeuPheValProHisTyrAsnThrAsnAlaHisSerIleIleTyrAlaLeuArgGlyArg	423
QY	1201	GCTCAGCTGCAAGTCTGTCAGACGCAACGCAACAGAGTGTACGACGAGGAGCTTCAAGAG	1260
DB	424	AlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnGlu	443
QY	1261	GGTCACTGCTGTTGCTGCCACAGAACTTCGCCCTGCTGGAAGTCCACAGCGAGAAC	1320
DB	444	GlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerAspAsn	463
QY	1321	TTGCAATACGTCGCTTCAAGACACTCAAGCCGACATAGCCCACTCCCGCTGAA	1380
DB	464	PheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnLeuAlaGlyGlu	483
QY	1381	AACTCCGTCATAGATAAAGCTCCGCGAGGAGGTGTGCAAAATTCATATGGCCCTCAAAGG	1440
DB	484	AsnSerValIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuGlnArg	503
QY	1441	GAGCAGGCAAGG---CAGCTTAAAGAACAAACCCCTTCAAGTTCCTCCACCGTCT	1497
DB	504	GluGlnAlaArgGlnGlnLeuLysAsnAsnAsnProPheLysPhePheValProProSer	523
QY	1498	CAGCAGTCTCCGAGGCTGTGCT 1521	
DB	524	GlnGlnSerProArgAlaValAla 531	

RESULT 3

Q9FZ11_ARAHY PRELIMINARY; PRT: 529 AA.

AC Q9FZ11; 01-MAR-2001 (TRENBLrel. 16, Created)

DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)

DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)

DE Gly1.

GN Name=Gly1;

OS Arachis hypogaea (Peanut).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;

OC Arachis.

OX NCBI_taxID=3818;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA	Jain A.K., Basha S.M.;
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF125192; AAG01363.1; -, mRNA.
DR	HSSP; P04776; IUC.
DR	GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR	InterPro; IPR006045; Cupin.
DR	InterPro; IPR007113; Cupin region.
DR	InterPro; IPR006044; Seedstore_11s.
DR	Pfam; PF00190; Cupin_1; 2.
DR	PRINTS; PR00439; IISGLOBULIN.
DR	PROSITE; PS00305; IIS_SEED_STORAGE; UNKNOWN_1.
SQ	SEQUENCE 529 AA; 60448 MW; BB6F25BC1D6E06A1 CRC64;

Alignment Scores:

Pred. No.: 1.2e-164 Length: 529

Score: 2479.00 Matches: 475

Percent Similarity: 95.3% Conservative: 9

Best Local Similarity: 93.5%		Mismatches: 20
Query Match: 89.2%		Indels: 4
DB:	2	Gaps: 2
US-10-728-323-3 (1-1524) x Q9FZ11_ARAHY (1-529)		
QY	1	CGGCAGCAACCGGAGGAGACCGCTGCCAGTTCCAGCGGCTCAATCGCGAGACCTGAC 60
DB	25	ArgGlnGlnProGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 44
QY	61	AATCGCATTAATCAGAGCGCGTTTACATTGAGACTTGGAAACCCCAACACAGGAGTTTC 120
DB	45	AsnArgLeuGluSerGluGlyGlyTyrIleGluThrTrpAsnProAsnAsnGlnGluPhe 64
QY	121	GAATGCGCGGCTCGCCCTCTCTCGCTTAGTCTCCGCGCAACCCCTTCGTAGGCTT 180
DB	65	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgArgPro 84
QY	181	TTCTACTCCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGGATACTTTGGTTG 240
DB	85	PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 104
QY	241	ATATTCCCTGTTGTCTTAGACACTATGAAGACCTCACACAAGGTCTTCATCTCAG 300
DB	105	IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgGlnHisGln 124
QY	301	TCCAAAGACCAACAAGAGCTCTTCAAGGAGAGACCAAGC---CAACAGCAACGAGAT 357
DB	125	SerGlnArgAlaProArgPheGluGlyGluAspGlnSerGlnGlnGlnArg 144
QY	358	AGTCACAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCAGTCCCAACCGGTGT 417
DB	145	SerHisGlnLysValArgArgPheAspGluGlyAspLeuIleAlaValProThrGlyVal 164
QY	418	GCCTTTCTGGCTCACAAACACACACGACTGATGTTGTTGCTCTTCTTACTGACAC 477
DB	165	AlaLeuTrpMetPheAsnAspHisAspThrAspValValAlaValSerLeuThrAspThr 184
QY	478	AACAACAACCAACAGCTTGATCATGTTCCCGAGAGATTCAATTTGGCTGGGACACG 537
DB	185	AsnAsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnHis 204
QY	538	GACCAAGCTTCTTAGGTACCAAGCAACAAGCAGACCAAGCAGACGAGAGACTTACCA 597
DB	205	GluGlnGluPheLeuArgTyrGlnGln-----GlnSerArgArgSerLeuPro 221
QY	598	TATAGCCCATACAGCCGCAAGCTCAGCCCTAGACAAGAGAGCGTCAATTTAGCCCTCGA 657
DB	222	TyrSerProTyrSerProGlnSerGlnProArgGlnGluArgGluPheSerProArg 241
QY	658	GGACAGCACAGCCGACAGAACGAGGAGCAAGAAGAAAGAAACCAAGGTGGAACATC 717
DB	242	GlyGlnHisSerArgArgGluArgAlaGlyGlnGluGluAsnGluGlyGlyAsnIle 261
QY	718	TTCAGCGGCTTACGCGGAGTTCCTGGACACAGCTTCAGGTTCACGACACAGATA 777
DB	262	PheSerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspAspArgGlnIle 281
QY	778	GTCAAAACCTTAAGAGCGAGACGAGAGTGAAGAAGAGGAGCGCAATTTGTGACAGTACG 837
DB	282	ValGlnAsnLeuArgGlyGluAsnGluSerGluGluGluGlyAlaIleValThrValLys 301
QY	838	GGAGGCTCAGATCTTGTAGCCCGAGATAGAAAAGACGTGCCACCAAGAGAGGATAC 897
DB	302	GlyGlyLeuArgIleLeuSerProAspArgLysArgGlyAlaAspGluGluGluTyr 321
QY	898	GATGAGATCAATATGATACGATGAAGAGATAGAAAGCGTGCAGGGGGAACGAGGC 957
DB	322	AspGluAspGluTyrGluTyrAspGluAspArgArgArgGlyArgGlySerArgGly 341
QY	958	AGGGGAATGGTATTGAAGAGACGATCTGCACCGCAAGTGTCTAAAAAGACATTGGTAGA 1017
DB	342	ArgGlyAsnGlyIleGluGluThrIleCysThrAlaSerValLysLysAsnIleGlyArg 361

QY	1018	AACAGATCCCTGACATCTAGAACCCCTCAAGCTGGTTGCTCACTCAAAACTGCCAACGATCTC 1077
DB	362	AsnArgSerProAspIleTyrAsnProGlnAlaGlySerLeuLysThrAlaAsnAspLeu 381
QY	1078	AACTTCTTAATCTAGTGGCTTGGACCTAGTGTGTAATATGGAATCTCTACAGGAAT 1137
DB	382	AsnLeuLeuIleLeuArgTrpLeuGlyLeuSerAlaGluTyrGlyAsnLeuTyrArgAsn 401
QY	1138	GCATTGTTTGTGCTCACTTACAAACCAACGACACAGCATCATATATCGATTGAGGGA 1197
DB	402	AlaLeuPheValProHisTyrAsnThrAsnAlaHisSerIleIleTyrAlaLeuArgGly 421
QY	1198	CGGCTCAGCTCAGTCTGAGACAGCAACGCAACAGAGTGTACGACGAGGAGCTTCAA 1257
DB	422	ArgAlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGln 441
QY	1258	GAGGCTCAGCTGCTTGTGTGTCACAGAACTTCGCGCTCGCTGGAAGTCCACAGACGAG 1317
DB	442	GluGlyHisValLeuValValProGlnAsnPheAlaValAlaGlySerGlnSerAsp 461
QY	1318	AACTTTCGAATACGTGGCATTCAAGACAGACTCAAGCCCGACATAGCCCACTCGCGGT 1377
DB	462	AsnPheGluTyrValAlaPheLysThrAspSerArgProAsnIleAlaAsnPheAlaGly 481
QY	1378	GAATACTCGCTCATAGATAACCTCGGAGGAGGTGGTTGCAATTCATATGCGCTCCAA 1437
DB	482	GluAsnSerIleIleAspAsnLeuProGluValValAlaAsnSerTyrGlyLeuPro 501
QY	1438	AGGAGGAGGCAAGGCGAGCTTAAGAACAAACACCCCTTCAAGTCTTCTTCCACCGTCT 1497
DB	502	ArgGluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPhePheValProProSer 521
QY	1498	CAGCAGTCTCCGAGGCTGTGGCT 1521
DB	522	GlnGlnSerLeuArgAlaValAla 529
RESULT 4		
ID	Q647H4_ARAHY	PRELIMINARY; PRT; 536 AA.
AC	Q647H4;	
DC	25-OCT-2004 (TrEMBLrel. 28, Created)	
DT	25-OCT-2004 (TrEMBLrel. 28, Last sequence update)	
DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)	
DE	Arachin Ahy-1.	
OS	Arachis hypogaea (Peanut).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;	
OC	rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;	
OC	Arachis.	
OX	NCBI_TaxID=3818;	
RN	[1]	
RP	NUCLEOTIDE SEQUENCE.	
RA	Yan Y., Wang L., Zhang Y.S., Wang L., Wu K., Huang S.Z.;	
RT	"Isolation of peanut genes encoding arachins and conglutins by	
RT	expressed sequence tags";	
RL	Plant Sci. 169:439-445(2005).	
RN	[2]	
RP	NUCLEOTIDE SEQUENCE.	
RA	Yan Y., Wang L., Huang S.;	
RT	"cDNA clone of peanut seed storage protein gene.";	
RL	Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; A772685; AAU21490.1; -; mRNA.	
DR	GO; GO:0045735; F:nutrient reservoir activity; IEA.	
DR	InterPro; IPR006045; Cupin.	
DR	InterPro; IPR007113; Cupin region.	
DR	InterPro; IPR006044; Seedstore_11s.	
DR	Pfam; PF00190; Cupin_1; 2.	
DR	PRINTS; PR00439; 11SGLBULIN.	
DR	PROSITE; PS00305; 11S SEED STORAGE; UNKNOWN_1.	
SQ	SEQUENCE 536 AA; 61506 MW; 30BDF638719AEB78 CRC64;	
Alignment Scores:		

SQ SEQUENCE 537 AA; 61532 MW; 3BDD034DEA159657 CRC64;

Alignment Scores:

Pred. No.: 1.01e-162 Length: 537
Score: 2451.50 Matches: 474
Percent Similarity: 93.6% Conservative: 10
Best Local Similarity: 91.7% Mismatches: 20
Query Match: 88.2% Indels: 13
DB: 2 Gaps: 4

US-10-728-323-3 (1-1524) x Q647H3_ARAHY (1-537)

```
QY 1 CGGCAGCAACCGGAGAGACGCGTCCAGTTCAGCGCCTCAATCGCAGACCTGAC 60
DB 24 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
QY 61 AATCGCATTAATCAGAGCGCGTTCATTGACACTTGGAAACCCCAACACGAGGAGTTC 120
DB 44 AsnArgLeuGluSerGluGlyGlyTyrIleGluThrTriPAsnProAsnAsnGlnPhe 63
QY 121 GAATGCGCGCGCTCGCCCTCTCTCGTTAGTCTCCGCGCGAACCGCCCTTCGTAGCCCT 180
DB 64 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro 83
QY 181 TTCCTACTCAAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGAGTACTTTGGGTG 240
DB 84 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlyArgGlyTyrPheGlyLeu 103
QY 241 ATATTCCTGTTGTCCTAGACACTATGAAGACGCTTCACACAAAGGTCGTGATCTCAG 300
DB 104 IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgArgHisGln 123
QY 301 TCCCAAGACCAACAAGACTCTTCCAAGGAGAGACAAAGC---CAACAGCAACGAGAT 357
DB 124 SerGlnArgAlaProArgPheGluGlyGluAspGlnSerGlnGlnGlnGlnAsp 143
QY 358 AGTCACAGAGGTGCACCGTTTCGATGAGGTGATCTCATTCAGTCCACCGGTGTT 417
DB 144 SerHisGlnLysValArgArgPheAspGluGlyAspLeuIleAlaValProThrGlyVal 163
QY 418 GCTTTCTGCTCTACACACACGACCTGATGATGTTGTTGCTGTTCTCTTACTGACACC 477
DB 164 AlaLeuTrpMetTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThr 183
QY 478 AACAAACACACACACAGCTTGATCTGTCCTCCAGGAGATTCAATTTGGCTGGGAACAG 537
DB 184 AsnAsnAsnAspAsnGlnLeuAspGlnPheProArgPheAsnLeuAlaGlyAsnHis 203
QY 538 GAGCAAGAGTTCCTTAAGGTACCAACCAACAAAGCAGACAAAGCAGAGAGAGCTTACCA 597
DB 204 GluGlnGluPheLeuArgTyrGlnGln-----GlnSerArgArgSerLeuPro 220
QY 598 TATAGCCCATACAGCCCGCAAGTTCAGCTAGACAAAGAGCGGTGAATTTAGCCCTCGA 657
DB 221 TyrSerProTyrSerProGlnSerGlnProArgGlnGluGluArgGluPheSerProArg 240
QY 658 GGACAGCACAGCCGACAGAACGACGACGACAGACAGACAAACGAGGTGAACATC 717
DB 241 GlyGlnHisSerArgArgGluArgAlaGlyGlnGlnGlnGluAsnGluGlyGlyAsnIle 260
QY 718 TTCAGCGGCTTCACGCGGAGTTCCTGGAAACAGCCCTCCAGGTTGACGACACAGATA 777
DB 261 PheSerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspAspArgGlnIle 280
QY 778 GTCAAAACCTTAAGCGGACGACCGGAGTGAAGAGAGGAGGAGCCATTGTGACAGTCAGG 837
DB 281 LeuGlnAsnLeuArgGlyGluAsnGluSerAspGluGlnGlyAlaIleValThrValArg 300
QY 838 GAGGCGCTCAGATCTTGACCCACGATAGAAAAGAG-----CGTGGCGAC 882
DB 301 GlyGlyLeuArgIleLeuSerProAspArgLysArgGlnGlnTyrGluArgProAsp 320
QY 883 GAAGAAGAGGAATACGATGAAGATGAATATGATATGATACGATCAAGAG-----GAT 930
```

```
DB 321 GluGluGluGluTyrAspGluAspGluTyrGluTyrAspGluGluGluArgGlnGlnAsp 340
QY 931 AGAAGCGCTGGCAGGGAGACGACGAGCGGGAATGTATTGAAGAGACGATCTGCACC 990
DB 341 ArgArgGlyArgGlySerArgGlyArgGlyAsnGlyIleGluGluThrIleCysThr 360
QY 991 GCAAGTCTAAAAGAACATTGGTAGAACACAGATCCCTGACATCTACAAACCTCAAGCT 1050
DB 361 AlaSerValLysLysAsnIleGlyArgAsnArgSerProAspIleTyrAsnProGlnAla 380
QY 1051 GGTTCCTCACTCAAACTGCCAACGATCTCAACCTTCTAATACTTAGTGGCTTCGACCTACT 1110
DB 381 GlySerLeuLysThrAlaAsnAspLeuAsnLeuLeuLeuArgTrpLeuGlyLeuSer 400
QY 1111 GCTGAATATGGAATCTCTACAGGAATGCAATGTTGTTGTCGCTCACTACAAACCAACGCA 1170
DB 401 AlaGluTyrGlyAsnLeuTyrArgAsnAlaLeuPheValProHisTyrAsnThrAsnAla 420
QY 1171 CACAGCATCATATATCGATTGAGGGACGGGCTCAGTGCACGCTCTTGTGTGCACAGCAACGC 1230
DB 421 HisSerIleIleTyrAlaLeuArgGlyArgAlaHisValGlnValValAspSerAsnGly 440
QY 1231 AACAGAGTGTACGACGAGGAGCTTCAAGAGGGTCACTGCTTGTGTGCACAGCAACTTC 1290
DB 441 AsnArgValTyrAspGluGluLeuGlnGlyHisValLeuValValProGlnAsnPhe 460
QY 1291 GCGGTGCTGGAAAGTCCAGAGCGAGAACTTCGAAATACGTGGCATCTCAAGACAGACTCA 1350
DB 461 AlaValAlaGlyLysSerGlnSerAspAsnPheGluTyrValAlaPheLysThrAspSer 480
QY 1351 AGCCCCAGCATAGCCCAACCTCCCGGTGAAACTCCGTATAGATAACCTCCGCGAGGAG 1410
DB 481 ArgProSerIleAlaAsnLeuAlaGlyGluAsnSerIleAspAsnLeuProGluGlu 500
QY 1411 GTGGTTGCAATTCATATGCTCCAAAGCGAGGAGGCAAGGAGCTTAAGAACCAAC 1470
DB 501 ValValAlaAsnSerTyrGlyLeuProArgGluGlnAlaArgGlnLeuLysAsnAsn 520
QY 1471 CCCTTCAAGTCTTCTGCTCCACCGTCTCAGCAGTCTCCGAGGGCTGTGGCT 1521
DB 521 ProPheLysPhePheValProProSerGlnGlnSerLeuGlyAlaValAla 537
```

RESULT 6

```
QBLKN1_ARAHY PRELIMINARY; PRT; 538 AA.
ID QBLKN1_ARAHY PRELIMINARY; PRT; 538 AA.
AC QBLKN1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Allergen Arah3/Arah4.
OS Arachis hypogaea (peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis
OC NCBI_TaxID=3818;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Viquez O.M., Konan K.N., Dodo H.W.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF510854; ASM46958.1; -; Genomic_DNA.
DR HSSP; P04776; IUC.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR pfam; PF00190; Cupin_1; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S_SEED_STORAGE; UNKNOWN 1.
SQ SEQUENCE 538 AA; 61737 MW; 7AABD0D59429709E CRC64;
```

Alignment Scores:

Pred. No.: 3,38e-162 Length: 538
Score: 2444.00 Matches: 471
Percent Similarity: 93.6% Conservative: 14
Best Local Similarity: 90.9% Mismatches: 19
Query Match: 87.9% Indels: 14
DB: 2 Gaps: 4

US-10-728-323-3 (1-1524) x Q8LKN1_ARAHY (1-538)

```
QY 1 CGGAGACCGGAGGAGGAGCGGTCAGTTCAGCGCCTCAATGCCGAGACCTGAC 60
DB 24 ArgGlnGlnProGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
QY 61 AATCGCATTGAATCAGAGCGCGTTCATTGAGACTTGGAAACCCCAACACCGAGGTTTC 120
DB 44 AsnArgIleGluSerGluGlyGlyTyrlleGluThrTrpAsnProAsnGlnGluPhe 63
QY 121 GAATGCCGCGGTCGCCCTCTCTCGTTAGTTCCTCCGCGCAACGCCCTTCGTAGGCT 180
DB 64 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgAsnAlaLeuArgPro 83
QY 181 TTCTACTCAATGCTCCCGAGAGATCTTCATCCAGGAGGAGGAGGATCTTGGTGTG 240
DB 84 PheTyrSerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgGlyTyrPheGlyLeu 103
QY 241 ATATTCCCTGGTGTCTAGACACTATGAAGAGCCTCACACAAGGTCGTGATCTCAG 300
DB 104 IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgHisGln 123
QY 301 TCCCAAGACCAACGAGCGTCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 124 SerGlnArgProProArgArgPheGlnGlyGlnAspGlnSerGlnGlnGlnAspSer 143
QY 361 CACCAAGGTCACCGTTTCATGAGGTCATCTCATTCAGTTCACCGGTCGTTGCT 420
DB 144 HisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla 163
QY 421 TTCTGCTCTACACGACCAACGACACTGATGTTGTGTGTTCTTCTTACTCACCAAC 480
DB 164 PheTyrMetTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 183
QY 481 AACACGACCAACGAGTTCAGTTCCTCCAGGAGATTCATTTGGCTGGGAGCAGGAG 540
DB 184 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu 203
QY 541 CAGAGTTTCTTAAAGTACAGCAACAAACGACGACGAGGAGGAGGAGGAGGAGGAGGAG 600
DB 204 GlnGluPheLeuArgTyrGlnGln-----GlnSerArgArgSerLeuProTyr 220
QY 601 AGCCCATACGCGCGCAAGTCCAGCTAGACAGGAGCGTGAATTTAGCCCTCGAGGA 660
DB 221 SerProTyrSerProGlnThrGlnProLysGlnGluAspArgGluPheSerProArgGly 240
QY 661 CAGCAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
DB 241 GlnHisGlyArgArgGluArgAlaGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 260
QY 721 AGCGGCTTCACCGCGGAGTTCCTGGAAACAGGCTTCAGGTTGACGACGACGAGTAGTG 780
DB 261 SerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspAspArgGlnIleLeu 280
QY 781 CAAACCTTAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
DB 281 GlnAsnLeuArgGlyGluAsnGluSerAspGluGlnGlyAlaIleValThrValArgGly 300
QY 841 GGCCTCAGATCTTGAGCCCGAGATGAGAAAGAGA-----CGTCCGACGAGAA 885
DB 301 GlyLeuArgIleLeuSerProAspArgLysArgGlnGlnTyrGluArgProAspGlu 320
QY 886 GAAGAGGAATACGATGAGATGAATATGATACGATGAAGAG-----GATAGA 933
DB 321 GluGluGluTyrAspGluAspGluTyrGluTyrAspGluGluGluArgGlnGlnAspArg 340
```

```
QY 934 AGCGGTGCGAGGGAAGCAGAGGCGGGAATATGTTATGAAGAGACGATCTGCACCCCA 993
DB 341 ArgArgGlyArgGlySerArgGlySerGlyAsnGlyIleGluGluThrIleCysThrAla 360
QY 994 AGTGTCTAAAAAACAATTGGTAGAAACAGATCCCTGACATCTACAAACCTCAAGCTGGT 1053
DB 361 SerPheLysLysAsnIleGlyArgAsnArgSerProAspIleTyrAsnProGlnAlaGly 380
QY 1054 TCACTCAAAACTGCCCAACGAT-----CTCAACCTTCTAATACTTAGTGGTGGCTGACCT 1107
DB 381 SerLeuLysThrAlaAsnGluLeuGlnLeuAsnLeuLeuLeuLeuArgTrpLeuGlyLeu 400
QY 1108 AGTGTGTAATATGGAATCTCTACAGGAATGCATTTTGTGCTCCTACACCAACCAAC 1167
DB 401 SerAlaGluTyrGlyAsnLeuTyrArgAsnAlaLeuPheValProHisTyrAsnThrAsn 420
QY 1168 GCACACGACATCATATATCGATTGAGGCGGCGGCTCAGCTGCAAGTCTGGACACGAC 1227
DB 421 AlaHisSerIleIleTyrAlaLeuArgGlyArgAlaHisValGlnValValAspSerAsn 440
QY 1228 GGCAACAGAGTGTACGACGAGGAGCTTCAAGAGGCTCACGTCGTTGTGTCGACAGAAC 1287
DB 441 GlyAspArgValPheAspGluGluLeuGlnGluGlyHisValLeuValValProGlnAsn 460
QY 1288 TTCGCGCTCGCTGGAACGTCCTCCAGAGCGAGAACTTCGAATACGTGGCATTAAGACAGAC 1347
DB 461 PheAlaValAlaGlyLysSerGlnSerGluAsnPheGluTyrValAlaPheLysThrAsp 480
QY 1348 TCAGGCGCCAGCATAGCCAACTCGCCGCTGAAACTCCGTCATAGATAAACCCTGCGGAG 1407
DB 481 SerArgProSerIleAlaAsnLeuAlaGlyGluAsnSerPheIleAsnLeuProGlu 500
QY 1408 GAGGTGGTTCGCAATTCATATGGCTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1467
DB 501 GluValValAlaAsnSerTyrGlyLeuProArgGluGlnAlaArgGlnLeuLysAsnAsn 520
QY 1468 AACCCCTTCAAGTTCCTCGTTCCACCGCTCTCAGCAGTCTCCGAGGCTGTGGCT 1521
DB 521 AsnProPheLysPhePheValProProSerGluGlnSerLeuArgAlaValAla 538
```

RESULT 7

Q6T2T4_ARAHY
ID Q6T2T4_ARAHY PRELIMINARY; PRT; 536 AA.
AC Q6T2T4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Storage protein.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RA Yang H.-X., Wang F., Bi Y.-P.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY439332; AAR02860.1; -, mRNA.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR006044; Seedstore_11a.
DR Pfam; PF00190; Cupin_1; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S_SEED_STORAGE; UNKNOWN 1.
SQ SEQUENCE 536 AA; 61498 MW; 138F3C8B1BC2C7D5 CRC64;

Alignment Scores:

Pred. No.:	1,22e-161	Length:	536
Score:	2436.00	Matches:	468
Percent Similarity:	93.6%	Conservative:	15
Best Local Similarity:	90.7%	Mismatches:	21

Query Match:	87.7%	Indels:	12
DB:	2	Gaps:	3
US-10-728-323-3 (1-1524) x Q6T2T4_ARAHY (1-536)			
QY	1	CGGCAGCAACCGGAGAGACGGTCCAGTTCACGCGCTCAATCGCGAGACCTGAC	60
DB	24	ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	43
QY	61	AATCGCATTAATCAGAGCGCGTTACATTGAGACTTGGAACCCCAACACGAGGATTC	120
DB	44	AsnArgIleGluSerGluGlyGlyTyrlleGluThrTrpAsnProAsnAsnGlnGluPhe	63
QY	121	GAATGGCGCGCGCTCGCCCTCTCTCGTTAGTCTCTCGCGCGCAACGCCCTTCGTAGCCCT	180
DB	64	GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgArgPro	83
QY	181	TTCTACTCCAACTGCCCGAGGAGATCTTCATCCAGCAAGGAGGGATCTTTGGTTG	240
DB	84	PheTySerAsnAlaProGlnGluIlePheIleGlnGlnGlyArgAlaTyrlleGlyLeu	103
QY	241	ATATTCCTCGTGTCTCTAGACACTATGAAGACCTCACACAAGGTCTGTCGATCTCAG	300
DB	104	IlePheLeuGlyCysProSerThrTyrlleGluProAlaGlnGlnGlyArgArgHisGln	123
QY	301	TCCCAAGACCAACAGACGCTCCAGAGGAGAACCAAGCCCAACAGCAACGAGATAGT	360
DB	124	SerGlnArgProProArgArgPheGlnGlyGlnAspGlnSerGlnGlnGlnAspSer	143
QY	361	CACCAGAGGTGCACCGTTTCGATGAGGGTGATCTCATTCAGTTCGCCCGGTGTGCT	420
DB	144	HisGlnIysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGlyValAla	163
QY	421	TTCTGCGCTTACACGACCAACGACCTGATGTGTGTGCTGTCTTCTTACTGACCAAC	480
DB	164	PheTrpMetTyrlleAsnAspHisAspThrAspValAlaValSerLeuThrAspThrAsn	183
QY	481	AACAACGACCAACAGCTTGATGATTCCTCCCGAGGAGATTCATTTGGCTGGGACACGG	540
DB	184	AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu	203
QY	541	CAAGAGTTCCTTAAGGTACCAAGCAACAAAGCAGACAAAGCAGACGAAAGAGACTTACCATAT	600
DB	204	GlnGluPheLeuArgTyrlleGlnGln-----GlnSerArgArgSerLeuProTyrlle	220
QY	601	AGCCCATACAGCCCGCAAGTTCAGCTAGACAAAGAGAGCGTGAATTTAGCCCTCGAGGA	660
DB	221	SerProTyrlleSerProGlnThrGlnProLysGlnGluAspArgGluPheSerProArgGly	240
QY	661	CAGCACAGCCGACAGACAGCAGGACAGACAGACAGACAGACAGGTCGAAACATCTTC	720
DB	241	GlnHisGlyArgArgGluArgAlaGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnPhe	260
QY	721	AGCGGCTTACGCGGAGTTCTCGGAACAGCTTCAGGTTCCAGGTTCCAGCACACAGATAGTG	780
DB	261	SerGlyPheThrProGluPheLeuAlaGlnAlaPheGlnValAspAspArgGlnIleLeu	280
QY	781	CAAAACCTAAGAGCGGACCGGAGTGAAGAGAGAGGAGCCATTGTGACAGTGAAGGGA	840
DB	281	GlnAsnLeuArgGlyGluAsnGluSerAspGluGlnGlyAlaIleValThrValArgGly	300
QY	841	GGCTCAGATCTTGAGCCAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG	885
DB	301	GlyLeuArgIleLeuSerProAspArgLysArgGlnGlnTyrlleGluArgProAspGlu	320
QY	886	GAAGAGAAATACATGAAGATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT	933
DB	321	GluGluGluTyrlleAspGluAspGluTyrlleGluTyrlleAspGluGluGluGlnGlnAspArg	340
QY	934	AGCGGTGGCAGGGAACGAGGAGGCGGGAATGGTATTGAAGAGACGATCTCGACCGCA	993
DB	341	ArgArgGlyArgGlySerArgGlySerGlyAsnGlyIleGluGluThrIleCysThrAla	360

QY	994	AGTGTCTAAAAGAACATTGGTAGAACAGATCCCTCATCATCAACCCCTCAAGCTGCT	1053
DB	361	SerPheLysLysAsnIleGlyArgAsnArgSerProAspIleTyrlleAsnProGlnAlaGly	380
QY	1054	TCACCTCAAAACTGCCCAACGATCTCAACCTTCTTAATCTAGTGGCTTTGGACCTAGTGT	1113
DB	381	SerLeuLysThrAlaAsnGluLeuAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	400
QY	1114	GAATATGGAATCTCTACAGGAATCGATTGTTGTCGCTCACTACACACCAACGACAC	1173
DB	401	GluTyrlleGlyAsnLeuTyrlleArgAsnAlaLeuPheValProHisTyrlleAsnThrAsn	420
QY	1174	ACATCATATATCGATTGAGGGACGGCTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG	1233
DB	421	SerIleIleTyrlleAlaLeuArgGlyArgAlaHisValGlnValValAspSerAsnGlyAsp	440
QY	1234	AGAGTGTACGACGAGGAGCTTCAAGAGGCTTCAAGAGGCTTCAAGTGTGTGTGTGTGTGT	1293
DB	441	ArgValPheAspGluGluLeuGlnGluGlyHisValLeuValValProGlnAsnPheAla	460
QY	1294	GTCGCTGGAAGTCCAGAGCGAGAACTTCGAATACGTCGTCGTCGTCGTCGTCGTCGTCG	1353
DB	461	ValAlaGlyLysSerGlnSerGluAsnPheGluTyrlleValAlaPheLysThrAspSerArg	480
QY	1354	CCAGCATAGCAACCTCCCGGTGAAACTCCGTCATAGATAACTCCGCGGAGGAGGTG	1413
DB	481	ProSerIleAlaAsnLeuAlaGlyGluAsnSerPheIleAspAsnLeuProGluVal	500
QY	1414	GTTGCAAAATTCATATGGCTCCAAAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1473
DB	501	ValAlaAsnSerTyrlleGlyLeuProArgGluGlnAlaArgGlnLeuLysAsnAsnPro	520
QY	1474	TTCAAGTTCCTGCTCCACCTCTCAGCAGTCTCCGAGGCTCTCGGCTGTGGCT	1521
DB	521	PheLysPhePheValProProSerGluGlnSerLeuArgAlaValAla	536
RESULT 8			
Q9SQH7_ARAHY	PRELIMINARY;	PRT;	530 AA.
ID	Q9SQH7;		
AC	Q9SQH7;		
DT	01-MAY-2000 (TrEMBLrel. 13, Created)		
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)		
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)		
DE	Glycinin.		
GN	Name=Arachis;		
OS	Arachis hypogaea (Peanut).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;		
OC	rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;		
OC	Arachis.		
OX	NCBI_TaxID=3818;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=Virginia;		
RX	MEDLINE=99406463; PubMed=10474031; DOI=10.1159/000024203;		
RA	Kleber-Janke T.; Cramer R.; Appenzeller U.; Schlaak M.; Becker W.-M.;		
RT	"Selective cloning of peanut allergens, including profilin and 2S		
RL	albumins, by phage display technology.";		
RL	Int. Arch. Allergy Immunol. 119:265-274 (1999).		
DR	EMBL; AF086821; AAD47382.1; -; mRNA.		
DR	HSSP; P04776; IUCX.		
DR	GO; GO:0045735; F.nutrient reservoir activity; IEA.		
DR	InterPro; IPR006045; Cupin.		
DR	InterPro; IPR007113; Cupin_region.		
DR	InterPro; IPR006044; Seedstore_11s.		
DR	Pfam; PF00190; Cupin_1; 2.		
DR	PRINTS; PR00439; 11SGLOBULIN.		
SQ	SEQUENCE 530 AA; 61011 MW; 864E349A31642A32 CRC64;		
Alignment Scores:			
Pred. No.:	1.36e-160	Length:	530
Score:	2421.00	Matches:	469
Percent Similarity:	94.1%	Conservative:	9

Best Local Similarity: 92.3% Mismatches: 29
 Query Match: 87.1% Indels: 2
 DB: 2 Gaps: 0

US-10-728-323-3 (1-1524) x Q9SQH7_ARAHY (1-530)

QY 1 CGGAGCAACCGGAGAGAACGGTGCAGTTCACGCGCTCAATGCGGAGAGACTGAC 60
 DB 24 ArgGlnGlnProGluGluAsnAlaCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 43
 QY 61 AATCCGATTGAATCAGAGCGCGGTACATTGAGACTTGGAACCCCAACACAGGAGTTC 120
 DB 44 AsnArgileGluSerGluGlyGlyrifleGluThrTrpAsnProAsnAsnGlnGluPhe 63
 QY 121 GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTTCCTCCGCGCAACGCGCTTCGTAGGCT 180
 DB 64 GluCysAlaGlyValAlaLeuSerArgLeuValLeuArgArgAsnAlaLeuArgPro 83
 QY 181 TTCTACTCCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGATACCTTTGGTGG 240
 DB 84 PheTyrSerAsnAlaProGlnGluPheIleGlnGlnGlyArgGlyTyrPheGlyLeu 103
 QY 241 ATATTCCTGGTGTCTCTAGACACTATGAGAGCCTCACACAAAGTGTGTCATCTCAG 300
 DB 104 IlePheProGlyCysProSerThrTyrGluGluProAlaGlnGlnGlyArgArgTyrGln 123
 QY 301 TCCCAAGACCAACGAGCTCTCAAGGAGAGAACCAAGCAACAGCAACGAGATAGT 360
 DB 124 SerGlnArgProProArgArgLeuGlnGluAspGlnSerGlnGlnGlnArgSer 143
 QY 361 CACAGAAAGTGACCGTTTCATGAGGTGATCTCATTCAGTTCCTCCACCGGTGTCT 420
 DB 144 HisGlnLysValHisArgPheAsnGluGlyAspLeuIleAlaValProThrGlyValAla 163
 QY 421 TTCTGCTCTACACGACACACACTGATGTGTCTGCTTCTCTACTACACCAAC 480
 DB 164 PheTrpLeuTyrAsnAspHisAspThrAspValValAlaValSerLeuThrAspThrAsn 183
 QY 481 AACACGACCAACGAGCTTCATGATTCCTCCAGAGATTCATTTGCTGGGAAACGAG 540
 DB 184 AsnAsnAspAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyAsnHisGlu 203
 QY 541 CAAGAGTTCTTAAGTTCACGACCAACAAACAGCAACAAAGCAGACGAAAGCTTACATAT 600
 DB 204 GlnGluPheLeuArgTyrGlnGlnGlnSerArgGlnSerArgArgSerLeuProTyr 223
 QY 601 ACCCATACAGCCCGCAAGTACGCTACACAGAGAGCGTGAATTTAGCCCTCGAGA 660
 DB 224 SerProTyrSerProHisSerArgProArgArgGluGluArgGluPheArgProArgGly 243
 QY 661 CAGCACAGCCGAGAGACGAGCAAGCAAGAAAGAAACCAAGGTGGAACATCTTC 720
 DB 244 GlnHisSerArgArgGluArgAlaGlyGlnGlnGluGluAspGluGlyGlyAsnIlePhe 263
 QY 721 AGCGGCTTCACCGGAGTTCCTGGAACAAAGCTTCAGTTCACGACGACAGATAGTG 780
 DB 264 SerGlyPheThrProGluPheLeuGluGlnAlaPheGlnValAspAspArgGlnIleVal 283
 QY 781 CAARACCTAAGGCGGACGAGTGAAGAGAGGAGCGCATTTGACAGTGAAGGGA 840
 DB 284 GlnAsnLeuTrpGlyGluAsnGluSerGluGluGluGlyAlaIleValThrValArgGly 303
 QY 841 GCGCTCAGATCTTGAAGCCAGATAGAAAGACGCTGCCGACGAAAGAGAGCAATACGAT 900
 DB 304 GlyLeuArgIleLeuSerProAspGlyThrArgGlyAlaAspGluGluGluGluThrAsp 323
 QY 901 GAAGATGAATATGAATACGATGAAGAGATAGAACGCTGGCAGGGGAGCAGAGCGAGG 960
 DB 324 GluAspGlnTyrGluTyrHisGluGlnAspGlyArgGlyArgGlySerArgGlyGly 343
 QY 961 GGGATGTTTGAAGAGAGCATCTGCACCGCAAGTGTCTAAAGAACATCTGTTGAAC 1020
 DB 344 GlyAsnGlyIleGluGluThrIleCysThrAlaCysValIleLysAsnIleGlyGlyAsn 363

QY 1021 AGATCCCTGCATCTACAACCCCTCAAGC-TGGTTTCACTCAAACTGCAACGATCTCAA 1079
 DB 364 ArgSerProHisIleTyrAspProGlnArgTrpPheThrGlnAsnCys-HisAspLeuAs 383
 QY 1080 CTTCTTAATACTAGGTGGCTTGGACCTAGTGTCTGAATATGGAATCTCTACAGGAATGC 1139
 DB 383 nLeuLeuIleLeuArgTrpLeuGlyLeuSerAlaGluTyrGlyAsnLeuTyrArgAsnAl 403
 QY 1140 ATTGTTTTCGCTCACTCAACACCAACGACACAGCATCATATATCGATTTCAGGGGACG 1199
 DB 403 aleuPheValProHisIleTyrAsnThrAsnAlaHisSerIleIleIleTyrAlaLeuArgGlyAr 423
 QY 1200 GGCTCACCTGCAAGTGTGGACGACCAACGCAACAGAGTGTACGACGAGGACTTCAAGA 1259
 DB 423 gAlaHisValGlnValValAspSerAsnGlyAsnArgValTyrAspGluGluLeuGlnG 443
 QY 1260 GGGTCACGTGTTGTGGTGCACAGAACTTCGCCCTCGTGGAAAGTCCACAGAGGAGAA 1319
 DB 443 uGlyHisValLeuValValProGlnAsnPheAlaValAlaGlyLysSerGlnSerGluAs 463
 QY 1320 CTTGCAATACGTGGCATTCAAGACAGACTCAAGGCCACAGCATAGCAACCTCGCGGTCA 1379
 DB 463 nPheGluTyrValAlaPheLysThrAspSerArgProSerIleAlaAsnPheAlaGlyG 483
 QY 1380 AAATCTCCGCTCATAGATAACCTCGCGGAGGAGTGTGTCAAATTCATATGGCTTCCAAAG 1439
 DB 483 uAsnSerPheIleAspAsnLeuProGluGluValValAlaAsnSerTyrGlyLeuProAr 503
 QY 1440 GGAGCAGCAGCAGCAGCTTAAGAACAAACACCCCTTCAAGTCTTCGTTCCACCGTCTCA 1499
 DB 503 gGluGlnAlaArgGlnLeuLysAsnAsnAsnProPheLysPheValProProPheG 523
 QY 1500 CGAGTCTCCGAGGCGTGGCT 1521
 DB 523 mGlnSerProArgAlaValAla 530

RESULT 9
 Q61NG5_ARAHY PRELIMINARY; PRT; 510 AA.
 ID Q61NG5; ARAHY PRELIMINARY; PRT; 510 AA.
 AC Q61NG5;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Glycinin (Fragment).
 OS Arachis hypogaea (Peanut).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
 OC Arachis.
 OX NCBI_TaxID=3818;
 RN (1)
 RP NUCLEOTIDE SEQUENCE.
 RA Kang I.-H., Gallo-Meagher M.;
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY618460; AAT39430.1; -, mRNA.
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
 DR InterPro; IPR006045; Cupin.
 DR InterPro; IPR006044; Seedstore_11s.
 DR Pfam; PF00190; Cupin 1; 2.
 DR PRINTS; PR00439; 11SGLOBULIN.
 DR PROSITE; PS00305; 11S_SEED_STORAGE; UNKNOWN_1.
 FT NON TER 1
 SQ SEQUENCE 510 AA; 58061 MW; F6569EF13D1A12BD CRC64;

Alignment Scores:
 Pred. No.: 1.45e-119 Length: 510
 Score: 1834.00 Matches: 364
 Percent Similarity: 78.6% Conservative: 43
 Best Local Similarity: 70.3% Mismatches: 71
 Query Match: 66.0% Indels: 40
 DB: 2 Gaps: 8

US-10-728-323-3 (1-1524) x Q6IWG5_ARAHY (1-510)

```
QY 1 CGCGACGAAACCGGAGAGAACGGCTGCCAGTTCACAGCGCTCAATCGGAGAGACCTGAC 60
Db |||||
QY 22 ArgGlnGlyGlyGluGluAsnGluCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp 41
Db |||||
QY 61 AATCGCATTAATCAGAGGGGGTACATTGAGACTTGGAAACCCCAACACGAGGAGTTC 120
Db |||||
QY 42 AsnArgIleGluSerGluGlyGlyTyrlleGluThrTrpAsnProAsnAsnGlnGluPhe 61
QY 121 GAATGGCGGGCGCTCGCCCTCTCTCGCTTAGTCTCCGCGGCAACCCCTTCGTAGCCCT 180
Db |||||
QY 62 GlnCysAlaGlyValAlaLeuSerArgThrValLeuArgArgAsnAlaLeuArgPro 81
QY 181 TTCTACTCCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGATGACTTTGGTTG 240
Db |||||
QY 82 PheTy-SerAsnAlaProLeuGluIleTyrlleValGlnGlnGlySerGlyTyrlleGlyLeu 101
QY 241 ATATTCCCTGGTTCCTAGACACTATGAGAGCCTTCACACAAAGTTCGTGATCTCAG 300
Db |||||
QY 102 IlePheProGlyCysProSerThrTyrlleGluGluProAlaGlnGluArgTyrlleGln 121
QY 301 TCCCAAGACCAACAGACGCTCCCAA-----GGAGAAAGACCAAGCCAAACAGCAACGA 354
Db |||||
QY 122 SerGlnLysProSerArgArgPheGlnValGlyGlnAspProSerGlnGlnGln 141
QY 355 GATAGTCACAGAGGTGCACCGTTTCGATGAGGGTGATCTCATTCAGTTCGCCACCGGT 414
Db |||||
QY 142 AspSerHisGlnLysValHisArgPheAspGluGlyAspLeuIleAlaValProThrGly 161
QY 415 GTTGCTTTCGGCTTACACGACCCAGCACATGATGTTGCTGCTTCTCTTACTGAC 474
Db |||||
QY 162 ValAlaPheTrpMetTyrlleAsnAspGluAspThrAspValThrValThrLeuSerAsp 181
QY 475 ACCAAACAAACACACACACGCTTGATTCAGTTCCTCCAGGAGATTCATTTGGCTGGGAAC 534
Db |||||
QY 182 ThrSerSerIleHisGlnLeuAspGlnPheProArgArgPheTyrlleAlaGlyAsn 201
QY 535 ACGGACGAAGTTCCTTAAGGTACCGAGCAACAAAGCAGACAAAGCAGACGAAGACTTA 594
Db |||||
QY 202 GlnGluGlnGluPheLeuArgTyrlleGlnGlnGlnGlySer----- 215
QY 595 CCATATAGCCCATACAGCCCGCAAGATCAGCTCCCTAGACAGAGCGGTGAATTTAGCCCT 654
Db |||||
QY 216 -----ArgProHisTyrlleArgGlnIleSerPro 224
QY 655 CGAGGACAGCAGCCCGCAGAGAACGAGCAGGAGCAAGAGAAACGAAGGTGGAAAC 714
Db |||||
QY 225 ArgValArg-----GlyAspGluGlnGluAsnGluGlySerAsn 237
QY 715 ATCTTCAGCGGCTTCACCGCGAGTTCCTGGAAACAGCCCTTCAGGTTGACGACAGACAG 774
Db |||||
QY 238 IlePheSerGlyPheAlaGlnGluPheLeuGlnHisAlaPheGlnVal---AspArgGln 256
QY 775 ATAGTCAAACCTTAAGAGCGGAGACCGCAGAGTGAAGAGAGGAGCCATTGTGACAGTG 834
Db |||||
QY 257 ThrValGlnAsnLeuArgGlyGluAsnGluArgGluGluGlnGlyAlaIleValThrVal 276
QY 835 AGGGGAGGCGCTCAGAATCTTCGAGCCCAAGATAGAGAGAGAGAGAGAGAGAGAGAGAG 879
Db |||||
QY 277 LysGlyGlyLeuArgIleLeuSerProAspGluGluAspGluSerArgSerArgProPro 296
QY 880 GACGAAGAAGGATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 939
Db |||||
QY 297 SerArgArgGluGluPheAspGluAspArg-----SerArgProGlnGlnArg 312
QY 940 GGCAGG-----GGAAGCAGAGGCGAGGGGAAATGCTATTCGAAGAGAGAGATCTGCACC 990
Db |||||
QY 313 GlyLysTyrlleAspGluAsnArgArgGlyTyrlleAsnGlyIleGluGlnThrIleCysSer 332
QY 991 GCAAGTCTAAAAGAAACATTGGTAGAAACAGATCCCTGACATCTACACACCTCAAGCT 1050
Db |||||
QY 333 AlaSerValLysLysAsnLeuLeuArgSerSerAsnProAspIleTyrlleAsnProGlnAla 352
```

```
QY 1051 GCTTCACTCAAACTGCCAAGCATCTCAACCTTCTAATACTTAGGTGGCTTGACCTAGT 1110
Db |||||
QY 353 GlySerLeuArgSerValAsnGluLeuAspLeuProIleLeuGlyTyrlleLeuGlyLeuSer 372
Db |||||
QY 1111 GCTGAATATGGAATCTCTACAGGAATGATTGTTGTCGCTCACTTACAAACACCAACGCA 1170
Db |||||
QY 373 AlaGlnHisGlyThrIleTyrlleArgAsnAlaMetPheValProHisTyrlleLeuAsnAla 392
Db |||||
QY 1171 CACGACATCATATATCGATTGAGGAGCGGCTCACTGTCGAAAGTCTGTGACACCAACGCGC 1230
Db |||||
QY 393 HisThrIleValValAlaLeuAsnGlyArgAlaHisValGlnValValAspSerAsnGly 412
QY 1231 AACAGAGTGTACAGCAGAGAGCTTCAAGAGGCTCAGCTTGTGGTGCACAGAACTTC 1290
Db |||||
QY 413 AsnArgValTyrlleAspGluGluLeuGlnGlyHisValLeuValValProGlnAsnPhe 432
QY 1291 GCGCTCGCTGGAAGTCCAGAGCAGAACTTCGAATACGTGGCATTTCAAGACAGACTCA 1350
Db |||||
QY 433 AlaValAlaAlaLysAlaGlnSerGluAsnTyrlleGluTyrlleAlaPheLysThrAspSer 452
QY 1351 AGGCCAGCATAGCCAACTCCCGCTGAAATCTCGTTCATAGATAAACCCTCGCGAGGAG 1410
Db |||||
QY 453 ArgProSerIleAlaAsnLeuAlaGlyGluAsnSerIleAlaAsnLeuProGluGlu 472
QY 1411 GTGGTTCGAATTCATATGCGCTCCCAAGGAGCAGGCAAGCAGCTTAAGAACACAAC 1470
Db |||||
QY 473 ValValAlaAsnSerTyrlleArgLeuProArgGluGlnAlaArgGlnLeuLysAsnAsn 492
QY 1471 CCCTTCAAGTTCCTGCTCCACCG--TCTCAGCAGTCTCCGAGGCTGTGGCT 1521
Db |||||
QY 493 ProPheLysPhePheValProPheAspHisGlnSerMetArgGluValAla 510
Db |||||
RESULT 10
Q647H2 ARAHY
ID Q647H2 ARAHY PRELIMINARY; PRT; 484 AA.
AC Q647H2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Arachis hly-3.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Yan Y.S., Lin X.D., Zhang Y.S., Wang L., Wu K., Huang S.Z.;
RT "Isolation of peanut genes encoding arachins and conglutins by
RT expressed sequence tags.";
RL Plant Sci. 169:439-445 (2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Yan Y., Wang L., Huang S.;
RT "cDNA clone of peanut seed storage protein gene.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY722687; AAU21492.1; -; mRNA.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR007113; Cupin region.
DR InterPro; IPR006044; Seedstore_1ls.
DR Pfam; PF00190; Cupin_1; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S SEED STORAGE; UNKNOWN.1.
SQ SEQUENCE 484 AA; 54569 MW; 5A3E950752E89D2D CRC64;
Alignment Scores: 8.9e-107 Length: 484
Pred. No.: 1651.00 Matches: 337
Score: 1651.00
Percent Similarity: 75.2% Conservative: 46
Best Local Similarity: 66.2% Mismatches: 76
```

Query Match:	59.4%	Indels:	50
DB:	2	Gaps:	8
US-10-728-323-3 (1-1524) x Q647H2_ARAHY (1-484)			
QY	1	CGGAGCAACCGGAGAGAACGGGTGCTCCAGGTTCAGCGCTCAATGCGGAGACCTGAC	60
DB	24	ArgGlnGlnGlyGluAsnGluCysGlnPheGlnArgLeuAsnAlaGlnArgProAsp	43
QY	61	AATCGCATTAATACAGAGCGCGGTACATTGAGACTTGGAACCCCAACCAACAGAGAGTTC	120
DB	44	AsnCysIleGlnSerGluGlyTyrlleGluThrTrpAsnProAsnAsnGlnGluPhe	63
QY	121	GAATGCGCGCGGTGCTCTCTCGCTTGTCTCTCCGCGCAACCGCTTCGTAGGCT	180
DB	64	GlnCysAlaGlyValAlaLeuSerArgPheValLeuArgArgAsnAlaLeuArgPro	83
QY	181	TTCTACTCAATGCTCCCGAGAGATCTTCATCCAGCAAGGAGGGGATACCTTTGGTGTG	240
DB	84	PheTyrSerAsnAlaProGlnGluIlePheIleTyrGlnGlySerGlyTyrPheGlyLeu	103
QY	241	ATATTCCCTGGTGTCTCTAGACACTATGAAGAGCCTCACACAAGGTGCTCGATCTCAG	300
DB	104	IlePheProGlyCysProGlyThrPheGluGluPro---IleGlnGly---SerGluGln	121
QY	301	TCCCAAGAGCCCAAGAGCGTCTCCAGGAGAGAACCAAGCAACAGCAACGAGATAGT	360
DB	122	PheGlnArgProSerArgHisPheGlnGlyGlnAspGlnSerGlnArgProLeuAspThr	141
QY	361	CACAGAGGTGACCGTTTCAGTAGAGGTGATCTCATTTGAGTTCCTCCACCGGTGTGCT	420
DB	142	HisGlnLysValHisGlyPheArgGluGlyAspLeuIleAlaValProHisGlyValAla	161
QY	421	TTCTGGCTCTACAACACCAACGACACTGATGTTGTCGTGTTCTTACTGACACCAAC	480
DB	162	TrpIleTyrAsnAspGlnAspThrAspValValAlaIleSerValLeuHisThrAsn	181
QY	481	AACAACGACAACCGAGTTCAGTTCCCGCAGGAGATTCAATTTGGTGGCAACGAGAG	540
DB	182	SerLeuHisAsnGlnLeuAspGlnPheProArgArgPheAsnLeuAlaGlyLysGlnGlu	201
QY	541	CAAGAGTTCCTTAAGGTACAGCAACAAGC---AGACAAGAGCAGCAAGAAGCTTACCA	597
DB	202	GlnGluPheLeuArgTyrGlnGlnArgSerGlyArgGlnSer-----	215
QY	598	TATAGCCCATACAGCCCGCAAGTACAGCTAGACAGAGAGAGCGTGCAATTAGCCCTCGA	657
DB	216	-----ProLysGlyGluGluGlnGlu-----	222
QY	658	GGACAGCACACGCGCAGAGAACGAGCAGGACAGAGAGAGAGAAACCAAGGTGGAACATC	717
DB	223	-----GlnGluGlnGluAsnGluGlyGlyAsnVal	232
QY	718	TTGAGCGCTTCACGCGGAGTTCCTGGAAACAGCCTTCAGGTTCAGCAGACAGAGATA	777
DB	233	PheSerGlyPheSerThrGluPheLeuSerHisGlyPheGlnValAsnGlu---AspIle	251
QY	778	GTCAAAACCTAAGAGCGCAGACCGAGAGTGTGAGAGAGGAGCGCCATTGTCAGCTGAGG	837
DB	252	ValArgAsnLeuArgGlyGluAsnGlnArgGluGluGlnGlyValAlaIleValThrValLys	271
QY	838	GGAGGCTTCAGATCTTGAGCCCATAGAAAGAGAGAGCGTCCGACGAAGAGGAGATAC	897
DB	272	GlyGlyLeuSerIleLeuValProPro-----	280
QY	898	GATGAAGATGAATATCAATACGATGAAGAGGATAGAAAGCGGTGGCAGGGGAACGAGAGC	957
DB	281	-----GluTrpArgGlnSerTyrGlnGlnProGlyArgGlyAspLysAsp	295
QY	958	AGGGGAATGGTATTGAAGAGAGATCTGACCGCAAGTGTCTAAAGAAGACATGGTAGA	1017
DB	296	PheAsnAsnGlyIleGluGluThrIleCysThrAlaThrValLysMetAsnIleGlyLys	315

QY	1018	AACAGATCCCTGTGACATCTACAAACCTCAAGCTGTGTTCACTCAAAACATGCCAACGATCTC	1077
DB	316	SerThrSerAlaAspIleTyrAsnProGlnAlaGlySerValArgThrValAsnGluLeu	335
QY	1078	RACCTTCTAAATCTTAGTGGCTTGGACCTAGTGTGAATATGGAATCTCTACAGGAAT	1137
DB	336	AspLeuProIleLeuAsnArgLeuGlyLeuSerAlaGluTyrGlySerIleHisArgAsp	355
QY	1138	GCATTGTTTGTGCTCACTACAAACCAACGACACAGCATCATATATCGATTGAGGGA	1197
DB	356	AlaMetPheValProHisTyrAsnMetAsnAlaAsnSerMetIleTyrAlaLeuHisGly	375
QY	1198	CGGCTCAGCTCAGTCAAGTGTGACAGCAACGCAACAGAGTGTACGACGAGAGCTTCAA	1257
DB	376	GlyAlaHisValGlnValValAspCysAsnGlyAsnArgValPheAspGluGluLeuGln	395
QY	1258	GAGGCTCAGCTGTGTCGCCACAGAACTTCGCGCTCGCTGGAAAGTCCCAGACGCGAG	1317
DB	396	GluGlyGlnSerLeuValValProGlnAsnPheAlaValAlaAlaLysSerGlnSerGlu	415
QY	1318	AACTTCGAATACGTGGCATTCGAAGACAGACTCAAGGCCAGCATAGCCAACTCGCCGCT	1377
DB	416	HisPheLeuTyrValAlaPheLysThrAsnSerArgAlaSerIleSerAsnLeuAlaGly	435
QY	1378	GAATACTCCGTCAATAGATAACCTCGCGAGAGGTGGTTCGAAATTCATATGCGCTCCAA	1437
DB	436	LysAsnSerTyrMetTrpAsnLeuProGluAspValValAlaAsnSerTyrGlyLeuGln	455
QY	1438	AGGAGCAGGAGCAGCTTAAGAACACCAACCCCTCAAGCTTCCTCCCTCCACCG---	1494
DB	456	TyrGluGlnAlaArgGlnLeuLysAsnAsnProPheThrPheLeuValProProGln	475
QY	1495	TCTCAGCAGCTCTCCGAGGCTGTGCT 1521	
DB	476	AspSerGlnMetIleArgThrValAla 484	
RESULT 11			
ID	GLY3_SOYBN	STANDARD;	PRT; 481 AA.
AC	P11828;		
DT	01-OCT-1989 (Rel. 12, Created)		
DT	01-OCT-1989 (Rel. 12, Last sequence update)		
DT	10-MAY-2005 (Rel. 47, Last annotation update)		
DE	Glycinin G3 precursor [Contains: Glycinin A subunit; Glycinin B subunit]		
GN	Name=GX3;		
OS	Glycine max (Soybean).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;		
OC	rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;		
OC	Glycine.		
OX	NCBI_TaxID=3847;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=cv. Dare; TISSUE=Leaf;		
RC	MEDLINE=89296500; PubMed=2740231;		
RA	Cho T.-J., Nielsen N.C.;		
RT	"The glycinin Gy3 gene from soybean."		
RL	Nucleic Acids Res. 17:4388-4388(1989).		
RN	[2]		
RP	DISCUSSION OF SEQUENCE.		
RX	MEDLINE=92393391; PubMed=2485233; DOI=10.1105/tpc.1.3.313;		
RA	Nielsen N.C., Dickinson C.D., Cho T.-J., Thanh V.H., Scallion B.J.,		
RA	Fischer R.L., Sims T.L., Drews G.N., Goldberg R.B.;		
RT	"Characterization of the glycinin gene family in soybean."		
CC	Plant Cell 1:313-328(1989).		
CC	!- FUNCTION: Glycinin is the major seed storage protein of soybean.		
CC	!- SUBUNIT: Hexamer; each subunit is composed of an acidic and a		
CC	basic chain derived from a single precursor and linked by a		
CC	disulfide bond.		
CC	!- SIMILARITY: Belongs to the 11S seed storage protein (globulins)		
CC	family.		

DE Glycinin G2 precursor [Contains: Glycinin A2 subunit; Glycinin B1a subunit].

DE Name=Gy2;

GN Glycine max (Soybean).

OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; OC Glycine.

OC NCBI_TaxID=3847;

RN [1] NUCLEOTIDE SEQUENCE.

RP STRAIN=cv. Dare; TISSUE=Leaf;

RC MEDLINE=92393391; PubMed=2485233; DOI=10.1105/tpc.1.3.313;

RX Nielsen N.C., Dickinson C.D., Cho T.-J., Thanh V.H., Scallion B.J., RA Fischer R.L., Sims T.L., Drews G.N., Goldberg R.B.;

RT "Characterization of the glycinin gene family in soybean.";

RL Plant Cell 1:313-328(1989).

RN [2] NUCLEOTIDE SEQUENCE.

RP STRAIN=cv. Dare; TISSUE=Leaf;

RC MEDLINE=89296499; PubMed=2740230;

RX Thanh V.H., Tumer N.E., Nielsen N.C.;

RA "The glycinin Gy2 gene from soybean.";

RT Nucleic Acids Res. 17:4387-4387(1989).

RN [3] NUCLEOTIDE SEQUENCE.

RP STRAIN=cv. Shirotsurunoko;

RC Utsumi S., Kim C.S., Kohno M., Kito M.;

RA "Polymorphism and expression of cDNAs encoding glycinin subunits.";

RT Agric. Biol. Chem. 51:3267-3273(1987).

RN [4] NUCLEOTIDE SEQUENCE.

RP MEDLINE=88040439; PubMed=3671077;

RX Fukazawa C., Momma T., Higuchi W., Uchida K.;

RA "Complete nucleotide sequence of the gene encoding a glycinin A2B1a subunit precursor of soybean.";

RT Nucleic Acids Res. 15:8117-8117(1987).

RN [5] NUCLEOTIDE SEQUENCE.

RP Momma T., Negoro T., Uchida K., Fukazawa C.;

RA "A complete cDNA coding for the sequence of glycinin A2B1a subunit precursor.";

RT FEBS Lett. 188:117-122(1985).

RN [6] PROTEIN SEQUENCE OF 19-296 AND 301-480 (A2 AND B1a SUBUNITS).

RP MEDLINE=85030470; PubMed=6541652;

RX Staswick P.E., Hermodson M.A., Nielsen N.C.;

RA "The amino acid sequence of the A2B1a subunit of glycinin.";

RT J. Biol. Chem. 259:13424-13430(1984).

RN [7] NUCLEOTIDE SEQUENCE OF 262-485.

RP MEDLINE=85030472; PubMed=6092376;

RX Marco Y.A., Thanh V.H., Tumer N.E., Scallion B.J., Nielsen N.C.;

RA "Cloning and structural analysis of DNA encoding an A2B1a subunit of glycinin.";

RT J. Biol. Chem. 259:13436-13441(1984).

RN [8] NUCLEOTIDE SEQUENCE OF 1-36.

RP MEDLINE=9032420; PubMed=2377465;

RX Kitamura Y., Arahira M., Itoh Y., Fukazawa C.;

RA "The complete nucleotide sequence of soybean glycinin A2B1a gene spanning to another glycinin gene Alab1b.";

RT Nucleic Acids Res. 18:4245-4245(1990).

RN [9] PROTEIN SEQUENCE OF 115-136 AND 202-224 (A2 SUBUNIT).

RP STRAIN=cv. Kishinevska-16;

RC MEDLINE=97054613; PubMed=8898910;

RX Shutov A.D., Kakhovskaya I.A., Bastrygina A.S., Bulmaga V.P., RA Horstmann C., Muntz K.;

RA "Limited proteolysis of beta-conglycinin and glycinin, the 7S and 11S storage globulins from soybean [Glycine max (L.) Merr.]. Structural and evolutionary implications.";

RT Eur. J. Biochem. 241:221-228(1996).

RN [10] DISULFIDE BOND.

RP MEDLINE=85030471; PubMed=6541653;

RX Staswick P.E., Hermodson M.A., Nielsen N.C.;

RA "Identification of the cysteines which link the acidic and basic components of the glycinin subunits.";

RT J. Biol. Chem. 259:13431-13435(1984).

RN [11] FUNCTION: Glycinin is the major seed storage protein of soybean.

CC -!- SUBUNIT: Hexamer; each subunit is composed of an acidic and a basic chain derived from a single precursor and linked by a disulfide bond.

CC -!- SIMILARITY: Belongs to the 11S seed storage protein (globulins) family.

CC -----

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC -----

CC EMBL; X15122; CAA33216.1; -; Genomic_DNA.

DR EMBL; D00216; BAA00154.1; -; mRNA.

DR EMBL; Y00398; CAA68460.1; -; Genomic_DNA.

DR EMBL; X02806; CAA26575.1; -; mRNA.

DR EMBL; X02646; AAA33963.1; -; Genomic_DNA.

DR EMBL; X53404; CAA37480.1; -; Genomic_DNA.

DR PIR; A91341; FWSYG1.

DR PIR; S11002; S11002.

DR HSP; P04776; 1FXZ.

DR SMR; P04405; 26-479.

DR InterPro; IPR006045; Cupin.

DR InterPro; IPR006044; Seedstore_11s.

DR Pfam; PF00190; Cupin_1; 2.

DR PRINTS; PR00439; 11SGLOBULIN.

DR PROSITE; PS00305; 11S_SEED_STORAGE; 1.

DR Direct protein sequencing; Multigene family; Polymorphism;

KW Seed storage protein; Signal; Storage protein.

KW -----

FT SIGNAL 1 18

FT CHAIN 19 296 Glycinin A2 subunit.

FT PROPEP 297 300

FT CHAIN 301 480 Glycinin B1a subunit.

FT PROPEP 481 485

FT DISULFID 104 307 Interchain (between A2 and B1a chains).

FT VARIANT 103 103 G -> D.

FT VARIANT 318 318 N -> T.

FT VARIANT 331 331 I -> V.

FT VARIANT 413 413 K -> R.

FT CONFLICT 39 39 D -> G (in Ref. 5).

FT CONFLICT 39 39 D -> N (in Ref. 6).

FT CONFLICT 61 61 C -> S (in Ref. 6).

FT CONFLICT 117 117 R -> C (in Ref. 6).

FT CONFLICT 343 343 W -> S (in Ref. 6).

SQ SEQUENCE 485 AA; 54391 MW; 78BB459837F77AD8 CRC64;

Alignment Scores:

Pred. No.:	3,46e-96	Length:	485
Score:	1499.50	Matches:	297
Percent Similarity:	72.0%	Conservative:	71
Best Local Similarity:	58.1%	Mismatches:	94
Query Match:	54.0%	Indels:	49
DB:	1	Gaps:	6

US-10-728-323-3 (1-1524) x GUYG2_SOYBN (1-485)

QY 1 CGGACGACCGGAGGAGAACGCGTCGAGCTTCAGCGCTCAATCGCAGACTGAC 60

DB 20 ArgGluGlnAlaGlnGlnAsnGluCysGlnLeuAsnAlaLeuLysProAsp 39

QY 61 AATCGCATTTGAATCAGAGGGGTTTACATTGAGACTTGAACCCCAACACGAGGAGTTC 120

DB 40 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 59

QY 121 GAATGCGCGCGCTCGCCCTCTCTCGCTTAGTCTCTCGCGCGCAACGCGCTTCGTAGGCTT 180

```
Db 60 GlnCysalaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 79
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
181 TTCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGATCTTTGGTTG 240
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 80 SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyAsnGlyIlePheGlyMet 99
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 241 ATATTCCTCGTGTCTAGACACTATGAAGACCTTCACACAAAGGTGTCGATCTCAG 300
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 100 IlePheProGlyCysProSerThrTyrGlnGluProGlnGluSerGlnGlnArgGlyArg 119
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 301 TCCCAAGACCAACCAAGACGCTCCCAAGGAGAGACCAAAAGCCAAACAGACATAGT 360
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 120 SerGlnArgPro-----GlnAspArg 126
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 361 CACCAGAAGGTGCACCGTTTCGATGAGGGTGATCTCATTCAGTCCCAACCGGTGTGCT 420
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 127 HisGlnLysValHisArgPheArgGluGlyAspLeuIleAlaValProThrGlyValAla 146
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 421 TTCTGCTCTACAAACGACACGACACTGATGTTGTTGCTGTTCTTCTTACTGACCAAC 480
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 147 TrpTrpMetTyrAsnAsnGluAspThrProValAlaValSerIleIleAspThrAsn 166
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 481 AACACGACAAACAGCTTGATCAGTCCCGAGGAGATTCAAATTGGCTGGACACGGAG 540
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 167 SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 541 CAAGAGTTCTTAAGGTACCAACCAACAAAGCAGACAAAGCAGACGAAAGCTTACCATAT 600
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 187 GlnGluPheLeuLysTyrGlnGln-----195
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 601 AGCCCATACAGCCCGCAAGTCAGCTAGACAGAGAGCGGTGAATTAGCTTCGAGGA 660
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 196 -----GlnGlnGlySerGlnSerGlnLysGly 205
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 661 CAGCACAGCCGAGAGAACGAGCAGGACCAAGAAAGAAACCAAGGTGGAACATCTTC 720
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 206 LysGln-----GlnGluGluAsnGluGlySerAsnIleLeu 218
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 721 AGCGGTTACGCGGAGTTCGTGGAACAGCTTCAGGTTCAGCAGACAGATAGTG 780
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 219 SerGlyPheAlaProGluPheLeuLysGluAlaPheGlyVal---AsnMetGlnIleVal 237
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 781 CAAACCTAAGACGCGACCGAGAGTGAAGAGAGGAGCGCCATTGTGACGTGAGGGA 840
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 238 ArgAsnLeuGlnGlyGluAsnGluGluAspSerGlyAlaIleValThrValLysGly 257
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 841 GGCCTCAGAACTTGAAGCCAGATAGAAAGAGACGTGCCGACGAAGAGGAATACGAT 900
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 258 GlyLeuArgValThrAlaProAlaMetArgLysProGlnGlnGluAspAspAsp 277
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 901 GAAGATGAA-----TATGAATACGATGAAGAGGATAGAAGCGGTGCGAGGGA 948
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 278 GluGluGluGlnProGlnCysValGluThrAspLysGlyCysGlnArgGlnSerLysArg 297
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 949 AGCAGAGGAGGGGATGTATTGAAGACATCTGCACCCCAAGTGTCTAAAAAGAAC 1008
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 298 SerArg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgGlnAsn 314
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 1009 ATTGGTAGAAACAGATCCCTGCATCTCAACACCTCAAGCTGGTTCACATCAAACTGCC 1068
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 315 IleGlyGlnAsnSerProAspIleTyrAsnProGlnAlaGlySerIleThrAla 334
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 1069 AAGCATCTCAACTTCTAATCTTAGTGGCTGGACCTAGTGTGAATATGAAATCTC 1128
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 335 ThrSerLeuAspPheProAlaLeuTrpLeuLeuLysLeuSerAlaGlnTyrGlySerLeu 354
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 1129 TACAGGAATCATGTTGTTCGCTCACTCAACACCAACGACACACGATCATATATCGA 1188
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 355 ArgLysAsnAlaMetPheValProHisTyrThrLeuAsnAlaAsnSerIleIleTyrAla 374
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Qy 1189 TTGAGGGGACGGGCTCACGTGCAAGTCGTGGACAGCAACCGCAACGAGGTGTACGACGAG 1248
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
```

```
Db 375 LeuAsnGlyArgAlaLeuValGlnValValAsnCysAsnGlyGluArgValPheAspGly 394
Qy 1249 GAGCTTCAAGAGGTCACGCTGTTGGTGGCCACAGAACTTCCCGTCCCTGGAAGTCC 1308
Db 395 GluLeuGlnGluGlyValLeuIleValProGlnAsnPheAlaValAlaAlaLysSer 414
Qy 1309 CAGAGCGAGAACTTCGAATACGTCGATTCAAGACAGACTCAAGGCCAGCATAGCCAAC 1368
Db 415 GlnSerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsn 434
Qy 1369 CTCGCGGTGAAACTCCGTCATAGATAACCTGCCGAGAGGTGGTTCGAAATTCATAT 1428
Db 435 LeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnHisThrPhe 454
Qy 1429 GGCCTTCAAGAGGAGCAGCAGGAGGCTTAAGAACAAACCCCTTCAAGTTCCTCGTT 1488
Db 455 AsnLeuLysSerGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuVal 474
Qy 1489 CCACCGCTCTCAGCAGTCTCCGAGGGCTGTGGCT 1521
Db 475 ProProGlnGluSerGlnArgArgAlaValAla 485

RESULT 13
Q549Z4 SOYBN
ID Q549Z4_SOYBN PRELIMINARY; PRT; 485 AA.
AC Q549Z4_2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Proglycinin A2B1.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
OC Glycine.
OC [1]_TaxID=3847;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Urade R., Nakatani H.;
RT "mRNA of Soybean Proglycinin A2B1 Subunit.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB113350; BAC78523.1; -; mRNA.
SQ SEQUENCE 485 AA; 54390 MW; 78BB459837F77AD8 CRC64;

Alignment Scores:
Pred. No.: 3,46e-96 Length: 485
Score: 1499.50 Matches: 297
Percent Similarity: 72.0% Conservative: 71
Best Local Similarity: 58.1% Mismatches: 94
Query Match: 54.0% Indels: 49
DB: 2

US-10-728-323-3 (1-1524) x Q549Z4_SOYBN (1-485)
Qy 1 CGSGACAACCCGAGGAGAACGCGTCAGGCTCCAGGCTCAATGCGCAGACCTGAC 60
Db 20 ArgGluGlnAlaGlnGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp 39
Qy 61 AATCGCATTAATCAGAGCGCGGTTCATTGAGACTTGAAACCCCAACACGAGGATTC 120
Db 40 AsnArgIleGluSerGluGlyGlyPheIleGluThrTrpAsnProAsnAsnLysProPhe 59
Qy 121 GAATGCGCGCGGTGCGCCCTCTCTCGTTAGTCTCCGCGCGCAACGCCCTCGTAGGCCT 180
Db 60 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgArgPro 79
Qy 181 TTCTACTCAATGCTCCCGAGGAGATCTTCATCCAGCAAGGAGGATCTTTGGTTG 240
Db 80 SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlyAsnGlyIlePheGlyMet 99
Qy 241 ATATTCCTCGTGTCTAGACACTATGAAGACCTTCACACAAAGGTGTCGATCTCAG 300
Db 100 IlePheProGlyCysProSerThrTyrGlnGluProGlnGluSerGlnGlnArgGlyArg 119
```

```
QY 301 TCCAAAGACCACCAAGACGCTCTCCAGGAGAGACCAAGCAACAGCAAGCAGATAGT 360
Db 120 SerGlnArgPro-----GlnAspArg 126
QY 361 CACCAGAGGTGCACCGCTTCGATGAGGTGATCTCATTTGCCAGTTCCACCGGTGTGTCT 420
Db 127 HisGlnLysValHisArgPheArgGluGlyAspLeuIleAlaValProThrGlyValAla 146
QY 421 TTCGTGCTCTACAAACACCAAGCAGACTGATGTTGTTGCTGTTCTCTTACTCACCAAC 480
Db 147 TrpTrpMetTyrAsnAsnGlnAspThrProValValAlaValSerIleIleAspThrAsn 166
QY 481 AACACGACCAACAGCTTGATCAGTTCCCGAGGAGATTCAATTTGGCTCGGACACGGAG 540
Db 167 SerLeuGluAsnGlnLeuAspGlnMetProArgArgPheTyrLeuAlaGlyAsnGlnGlu 186
QY 541 CAGAGTTCTTAAAGTTACAGCAACCAAGCAGCAACCAAGCAGCAGCAAGAGCTTACCAT 600
Db 187 GlnGluPheLeuLysTyrGlnGlnGln----- 195
QY 601 AGCCCATACAGCCCGCAAGTCAGCTCAGCAGCAAGAGAGCGTGAAATTTAGCCTCGAGA 660
Db 196 -----GlnGlnGlySerGlnSerGlnLysGly 205
QY 661 CAGCACAGCCGACAGAGACGAGCAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 206 LysGln-----GlnGluGluGluAsnGluGlySerAsnIleLeu 218
QY 721 AGCGGCTTACGCGCGAGTTCCTGGAAACAAGCCTTCAGGTTGACGACAGCAGATAGTG 780
Db 219 SerGlyPheAlaProGluPheLeuLysGluAlaPheGlyVal---AsnMetGlnIleVal 237
QY 781 CAAACCTTAAGAGCCGAGCCGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 238 ArgAsnLeuGlnGlyGluAsnGluGluAspSerGlyAlaIleValThrValLysGly 257
QY 841 GGCCTCAGATCTTGAGCCGACAGTAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 258 GlyLeuArgValThrAlaProAlaMetArgLysProGlnGlnGluGluAspAspAsp 277
QY 901 GAAGATGAA-----TATGAATACGATGAAGAGAGATGAAGAGAGAGAGAGAGAGAG 948
Db 278 GluGluGluGlnProGlnCysValGluThrAspLysGlyCysGlnArgGlnSerLysArg 297
QY 949 AGCAGAGCGAGGGGAATGTTTGAAGAGAGAGATCTGCACCGCAAGTCTAAAGAGAC 1008
Db 298 SerArg-----AsnGlyIleAspGluThrIleCysThrMetArgLeuArgGlnAsn 314
QY 1009 ATTGGTAGAAACAGATCCCTGACATCTACACCTCAAGCTGAGTTCCTCAAAAGTCC 1068
Db 315 IleGlyGlnAsnSerSerProAspIleTyrAsnProGlnAlaGlySerIleThrAla 334
QY 1069 AAGCATCTCAACCTTCTAATCTTAGTGCTGGACCTAGTGTGCAATATGAATCTC 1128
Db 335 ThrSerLeuAspPheProAlaLeuTrpLeuLeuLysLeuSerAlaGlnTyrGlySerLeu 354
QY 1129 TACAGGAATGATGTTGTCGCTCACTACCAACCAACGACAGCATCATATATCGA 1188
Db 355 ArgLysAsnAlaMetPheValProHisTyrThrLeuAsnAlaAsnSerIleIleTyrAla 374
QY 1189 TTGAGGGGACGGGCTCACGTGCAAGTCGTGGACAGCAACGCGCAACAGAGTGTACGACGAG 1248
Db 375 LeuAsnGlyArgAlaLeuValGlnValValAsnCysAsnGlnGluArgValPheAspGly 394
QY 1249 GAGCTTCAAGAGGTCAAGTCTGTTGGTGGCCACAGAACTTCGCGCTCGCTGGAAGTCC 1308
Db 395 GluLeuGlnGluGlyGlyValLeuIleValProGlnAsnPheAlaValAlaLysSer 414
QY 1309 CAGAGCGAGAACTTCCAACTAGTGGCATTCAAGACAGACTCAAGGCCACGACATAGCAAC 1368
Db 415 GlnSerAspAsnPheGluTyrValSerPheLysThrAsnAspArgProSerIleGlyAsn 434
```

```
QY 1369 CTCGCGGTGAAAACTCCGTCATAGATAACCTGCCGAGGAGGTGGTTCCAAATTCATAT 1428
Db 435 LeuAlaGlyAlaAsnSerLeuLeuAsnAlaLeuProGluGluValIleGlnHisThrPhe 454
QY 1429 GCGCTCCAAAGGAGGAGCAGCAGGAGCTTAGAACAACAACCCCTTCAAGTTCCTCGTT 1488
Db 455 AsnLeuLysSerGlnGlnAlaArgGlnValLysAsnAsnAsnProPheSerPheLeuVal 474
QY 1489 CCACGCTCTCAGCAGCTCTCCGAGGCTGTGGCT 1521
Db 475 ProProGlnGluSerGlnArgArgAlaValAla 485

RESULT 14
GLYGI_SOYBN
ID GLYGI_SOYBN STANDARD; PRT; 495 AA.
AC P04776;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Glycinin G1 precursor [Contains: Glycinin Ala subunit; Glycinin Bx subunit].
GN Name=GVI;
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
OC Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=cv. Dare; TISSUE=Leaf;
RX MEDLINE=92393391; PubMed=2485233; DOI=10.1105/tpc.1.3.313;
RA Nielsen N.C., Dickinson C.D., Cho T.-J., Thanh V.H., Scallion B.J.,
RA Fischer R.L., Sims T.L., Drews G.N., Goldberg R.B.;
RT "Characterization of the glycinin gene family in soybean.";
RL Plant Cell 1:313-328 (1989).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=cv. Dare; TISSUE=Leaf;
RX MEDLINE=89296498; PubMed=2740229;
RA Sims T.L., Goldberg R.B.;
RT "The glycinin Gyl gene from soybean.";
RL Nucleic Acids Res. 17:4386-4386 (1989).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=cv. Boninori;
RX MEDLINE=86041867; PubMed=2997720;
RA Negoro T., Momma T., Fukazawa C.;
RT "A cDNA clone encoding a glycinin Ala subunit precursor of soybean.";
RL Nucleic Acids Res. 13:6719-6731 (1985).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Utsuni S., Kohno M., Mori T., Kito M.;
RT "An alternate cDNA encoding glycinin Ala Bx subunit.";
RL J. Agric. Food Chem. 35:210-214 (1987).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RA Urade R., Nakatani H., Nakano C.;
RT "mRNA of soybean proglycinin AlaB1b subunit.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Glycinin is the major seed storage protein of soybean.
CC -!- SUBUNIT: Hexamer; each subunit is composed of an acidic and a
CC basic chain derived from a single precursor and linked by a
CC disulfide bond.
CC -!- PTM: The precursor is post-translationally processed to form a
CC covalently linked Ala-Bx subunit complex.
CC -!- SIMILARITY: Belongs to the 11S seed storage protein (globulins)
CC family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
```

```
CC removed.
CC -----
DR EMBL; M36686; AAA33966.1; -; mRNA.
DR EMBL; X15121; CAA33215.1; -; Genomic_DNA.
DR EMBL; X02985; CAA36723.1; -; mRNA.
DR EMBL; AB113349; BAC78522.1; -; mRNA.
DR PIR; A23497; FWSYG2.
DR PIR; S10851; S10851.
DR PDB; 1FXZ; X-ray; A/B/C=20-495.
DR PDB; 1UCX; X-ray; A/B/C=20-495.
DR PDB; 1UD1; X-ray; A/B/C=20-495.
DR InterPro; IPR006045; Cupin_region.
DR InterPro; IPR007113; Cupin_region.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin_1; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
DR PROSITE; PS00305; 11S_SEED_STORAGE; 1.
KW 3D-structure; Multigene family; Seed storage protein; Signal;
KW Storage protein.
FT SIGNAL 1 19 Glycinin Ala subunit.
FT CHAIN 20 306 Glycinin Bx subunit.
FT PROPEP 307 310 Glycinin Bx subunit.
FT CHAIN 311 490 Interchain (between Ala and Bx chains)
FT PROPEP 491 495 (By similarity).
FT DISULFID 107 317 D -> G (in Ref. 3).
FT CONFLICT 42 42 P -> S (in Ref. 3).
FT CONFLICT 108 108 F -> S (in Ref. 3).
FT CONFLICT 136 136 E -> G (in Ref. 3).
FT CONFLICT 360 360
FT TURN 30 31
FT STRAND 39 39
FT STRAND 43 47
FT TURN 48 49
FT STRAND 50 54
FT TURN 57 58
FT TURN 60 65
FT TURN 66 66
FT STRAND 68 74
FT TURN 76 77
FT STRAND 78 84
FT STRAND 89 95
FT STRAND 97 103
FT TURN 105 106
FT STRAND 134 137
FT TURN 138 139
FT STRAND 140 144
FT TURN 146 147
FT STRAND 149 154
FT STRAND 160 166
FT TURN 168 169
FT TURN 171 172
FT STRAND 181 183
FT TURN 191 192
FT TURN 193 196
FT TURN 222 223
FT TURN 226 233
FT TURN 234 234
FT TURN 237 243
FT STRAND 255 257
FT TURN 317 318
FT STRAND 322 324
FT STRAND 333 335
FT TURN 336 338
FT STRAND 339 344
FT TURN 346 348
FT TURN 350 352
FT TURN 353 356
FT STRAND 359 364
FT TURN 366 367
FT STRAND 369 375
FT TURN 376 377
FT STRAND 380 386
```

```
FT STRAND 388 394
FT TURN 396 397
FT STRAND 400 407
FT TURN 408 409
FT STRAND 411 414
FT TURN 416 417
FT STRAND 419 424
FT TURN 427 428
FT STRAND 429 435
FT STRAND 442 444
FT TURN 448 449
FT TURN 451 451
FT HELIX 452 454
FT HELIX 457 464
FT TURN 465 465
FT HELIX 468 476
FT STRAND 483 484
SQ SEQUENCE 495 AA; 55706 MW; 329CB0545B24D894 CRC64;

Alignment Scores:
Pred. No.: 6,1e-96 Length: 495
Score: 1496.00 Matches: 292
Percent Similarity: 70.9% Conservative: 73
Best Local Similarity: 56.7% Mismatches: 100
Query Match: 53.8% Indels: 50
DB: 1 Gaps: 6

US-10-728-323-3 (1-1524) x GLYGI_SOYBN (1-495)
QY 1 CGCGAGCAACCGGAGGAGAACGCGTCCAGTTCAGCGCTCAATGCGCAGAGACCTGAC 60
Db 23 ArgGluGlnProGlnGlnAsnGluCysGlnIleGlnLysLeuAsnAlaLeuLysProAsp 42
QY 61 AATCGCATTGAATCAGAGGCGGTTACATTGAGACTTGAACCCCAACCAACGAGGTTC 120
Db 43 AsnArgIleGluSerGluGlyGlyLeuIleGluThrTrpAsnProAsnAsnLysProPhe 62
QY 121 GAATGCGCGGCGTGCCTCTCTCGCTTAGTTCCTCCGCGCAACGCCCTTCGTAGGCCT 180
Db 63 GlnCysAlaGlyValAlaLeuSerArgCysThrLeuAsnArgAsnAlaLeuArgPro 82
QY 181 TTCTACTCCAATGCTCCCGAGGAGATCTTCATCCAGAGGAGGAGTACTTTGGTTG 240
Db 83 SerTyrThrAsnGlyProGlnGluIleTyrIleGlnGlnGlySerGlyIlePheGlyMet 102
QY 241 ATATTCCCTGGTTGCTTAGACACTATGAAGAGCCTCACACAAAGTTCGTGATCTCAG 300
Db 103 IleTyrProGlyCysProSerThrPheGluGluProGlnGlnProGlnGlnArgGlyGln 122
QY 301 TCCCAAGACCAACCAAGACGTCTCCNAGGAGAGCAACCAAGCAACGAGATAGT 360
Db 123 SerSerArgPro-----GlnAspArg 129
QY 361 CACCAAGGTGCACCGTTTCGATGAGGTGTATCTCATTCAGTTCACCGGTGTGCT 420
Db 130 HisGlnLysIleTyrAsnPheArgGluGlyAspLeuIleAlaValProThrGlyValAla 149
QY 421 TTCTGGCTCTACAACGACCAACGACACTGATGTTGTTGTTCTTCTTACTGACCAAC 480
Db 150 TrpTrpMetTyrAsnAsnGluAspThrProValValAlaValSerIleIleAspThrAsn 169
QY 481 AACACGACACACGAGCTTCAGTTCCTCCAGGAGATTCAATTGCTGGACACGGAG 540
Db 170 SerLeuGluAsnGlnLeuAspGlnMetProArgPheTyrLeuAlaGlyAsnGlnGlu 189
QY 541 CAAGAGTTCTTAAGGTACCAGCAACAAAGCAGACAAAGCAGACGAGAAAGCTTACCATAT 600
Db 190 GlnGluPheLeuLysTyrGlnGlnGluGln----- 199
QY 601 AGCCCATACAGCCCGCAAAAGTCAGCCTAGACAAAGAGAGCGTGAAATTAGCCCTCGAGGA 660
Db 200 -----GlyGly 201
```


